NORTH ATLANTIC TREATY ORGANIZATION
DEFENCE RESEARCH GROUP

TECHNICAL REPORT
AC/243(Panel 8)TR/1

A DIRECTORY OF HUMAN PERFORMANCE MODELS FOR SYSTEM DESIGN

Panel 8 on the Defence Applications of Human Performance Models for System Design

RSG.9 on Modelling of Human Operator Behaviour in Weapon Systems

CONSEIL DE L'ATLANTIQUE NORD
NORTH ATLANTIC COUNCIL

UNCLASSIFIED/UNLIMITED

ORIGINAL: ENGLISH
27 December 1991

TECHNICAL REPORT
AC/243(Panel 8)TR/1

DEFENCE RESEARCH GROUP

PANEL 8 ON THE DEFENCE APPLICATIONS OF HUMAN AND BIO-MEDICAL SCIENCES

Technical Report on the Study on A Directory of Human Performance Models for System Design

This is the Technical Report on the study on a Directory of Human Performance Models for System Design. It was prepared by RSG.9 on Modeling of Human Operator Behaviour in Weapon Systems. The Executive Summary of this report ("Yellow Pages") will be distributed under reference AC/243-N/346 dated 8 January 1992.

(Signed) Dr. J. VERMOREL
Defence Research Section

NATO,
1110 Brussels.

AC/243(P/8)TR/1
19200232

UNCLASSIFIED/UNLIMITED

REPORT DOCUMENTATION PAGE

1. Recipient's Reference:	2. Further Reference:
3. Originator's Reference: AC/243(Panel 8)TR/1	4. Security Classification: UNCLASSIFIED/UNLIMITED 5. Date: 27 DEC 91 — 6. Total Pages: 316 p.

7. Title (NU):
A DIRECTORY OF HUMAN PERFORMANCE MODELS FOR SYSTEM DESIGN

8. Presented at:

9. Author's/Editor's:
G.R. MC MILLAN, D. BEEVIS, W. STEIN, M.H. STRUB, E. SALAS, R. SUTTON, K.C. REYNOLDS

10. Author(s)/Editor(s) Address:
G.R. Mc Millan
Human Engineering Division
Armstrong Laboratory
Wright-Patterson AFB OH
45433-6573

11. NATO Staff Point of Contact:
Defence Research Section
NATO Headquarters
B-1110 Brussels
Belgium
(Not a Distribution Centre)

13. Keywords/Descriptors:
HUMAN PERFORMANCE, MODELS, SYSTEM DESIGN, WORKLOAD MODELS, BIOMECHANICAL MODELS, ANTHROPOMETRIC MODELS, TRAINING MODELS, NETWORK MODELS, DECISION MAKING MODELS, MANUAL CONTROL MODELS, CREW PERFORMANCE MODELS

14. Abstract:

This report catalogues over fifty human performance models and model development tools which are applicable to system design. It provides potential model users with brief reviews of each model, presented in a standard format. It is meant to be a practical source book, and does not address mathematical or theoretical issues to any great extent. Each chapter is aimed at a specific problem area in the system design/development cycle. Chapter topics include Task Allocation and Workload Prediction, Single Task Models, Multi-Task Models, Multi-Operator Models, Biomechanics and Work Space Design, Training and Skill Retention Models, and Network Modeling Tools.

TABLE OF CONTENTS

This page has been left blank intentionally.

CHAPTER 0

EXECUTIVE SUMMARY

0.1 SUMMARY OF THE STUDY

(i) Human factors engineering has evolved out of the early work in applied experimental psychology through the design handbook era and is becoming more interdisciplinary with strong technological influences from computer and information science, operations research, and systems simulation and modelling. Human factors/ergonomics models have the potential for representing human performance in ways which are compatible with both operations research and systems engineering. Despite this, human factors models are not widely used, either by human engineering specialists or system designers.

(ii) In response to this, Research Study Group 9 of DRG Panel 8 was convened to consolidate the available knowledge, to stimulate information exchange and cooperative research, to foster the practical application of modelling research, and to provide a bridge between models and approaches adopted by engineers and behavioural scientists.

(iii) Acting on these terms of reference, the RSG:

(1) reviewed current human performance models which are in use;

(2) investigated the development of micro-computer based models of human performance;

(3) conducted a technology demonstration workshop which included working demonstrations of typical models, and technical papers on their application

(iv) The presentations and discussions of the workshop were published in a book which provides an overview of the state of the art of human performance modelling (McMillan, G.R., Beevis, D., Salas, E., Strub, M., Sutton, R., van Breda, L. (Eds.) (1989). Applications of Human Performance Models to System Design. New York: Plenum). The investigation of micro-computer based models revealed that a number of models are being hosted on micro-computers, and suggested that such developments are likely to increase given government and user support. The review of current models identified 54 which were of sufficient interest to document in this report, in the form of a directory.

(v) Although this directory is provided to meet the first objective of the RSG, it is intended to serve broader aims by providing potential users with brief reviews of models reported in a standard format. Each major section of the report is aimed at a specific problem area in the system design/development cycle, based on the complementary knowledge and experience of the contributing nations and services. Model categories include Task Allocation and Workload Prediction, Single Task Models, Multi-Task Models, Multi-Operator Models, Biomechanics and Work Space Design, and Training and Skill Retention. A section of the directory also reviews tools which are available to support the development of new models, or the modification of existing models. In order to be included in the review, a model had to meet the criteria that it has the potential for solving a practical design

problem, and as a consequence, is not simply an interesting idea that has not been developed.

Table 0.1

Summary of Status of Models Reviewed

Model Type	Workload Prediction	Single Task	Multi-Task	Multi-Operator	Biomechanical & Work Space	Training & Skill Retention
Number Reviewed	7	16	10	6	9	6
Validated	2	9	4	3	4	4
Proprietary	4	2	4	2	4	0
Readily Available	2	9	4	2	3	6
Widely Used	1	3	2	1	2	3

(vi) Ten of the models reviewed appear applicable to the concept development stage of system design. By far, the majority of the models are applicable to later design stages. Most models require quantitative details that are not available at the concept development stage. Some of the training and skill retention models are not limited to application during system design. They are also applicable to deployed systems and current operations.

(vii) The review provided some understanding of why available models are not widely used. Table 0.1 summarises the information on availability, usage and validation of the 54 models. Only 26 of the models appear to have been validated in some way, usually by one or two limited experimental comparisons. Only 26 of the models can be classified as readily available (through publications and/or software). Many models appear to have been developed and used successfully in one or two projects, then ignored. These findings support one of the conclusions of a previous Panel 8 workshop that more effort needs to be put into the application of those human factors tools which have been developed, rather than undertaking the development of new tools or models without regard for their eventual application. To foster such applications, the RSG.9 technology demonstration workshop emphasised presentations on micro-computer hosted models which system designers could use.

(viii) Most of the models reviewed were "normative". That is, they represent an "ideal" operator, or are based on a theory of what "should" happen. This is appropriate for the design of new man-machine systems. Many of the models could be manipulated to represent "non-ideal" performance and, as such, they might be applicable in combat effectiveness simulations. Some are also related to human response to operational/battle stress and fatigue, which has recently become a subject of interest to Panel 7. Such models contain terms that permit one to represent the differences between "idealised" and "real" operators, the effects of environmental and battle stressors, the differences between ideal and combat effectiveness, or the differences between individual and group performance.

(ix) A major deficiency in the range of available models is that no reliable or validated models of cognitive tasks such as planning and decision making are available. This deficiency is a reflection of the immaturity of our understanding of cognitive task performance. Given the lack of appropriate analytic decision models, some modellers are using SAINT (Systems Analysis of Integrated Networks of Tasks) or SLAM (Simulation Language for Alternative Modeling) -based probabilistic network models in attempts to represent decision making in military systems. Until our understanding of human decision making is developed further, this is probably the only path available to the system modellers.

0.2 MAIN CONCLUSIONS

(x) Fifty-four models of human operator performance were found which are applicable to the design and development of weapon systems and their associated training systems.

(xi) These models apply to a variety of design activities, including Operator Workload Prediction, Single Task Performance, Multi-Task Performance, Multi-Operator Performance, Biomechanical Analysis and Workspace Design, and Training and Skill Retention.

(xii) No general-purpose model of operator performance exists; instead, system designers must select a model which is appropriate to the specific task, or tasks, which are being studied.

(xiii) One of the most effective ways of using such models is to reduce the range of potential design solutions to a manageable number. For example, workload models can be used to highlight the need for a simulator study of a task; control theory models can be used to select the critical conditions for evaluation; or anthropometrical models can be used to reduce the range of potential crew station configurations to the point where a less costly mockup can be constructed.

(xiv) One of the most promising areas of model development is in the use of tools such as SAINT to build network models of operator tasks. This approach, which typically uses Monte-Carlo simulation, is compatible with current trends in functional analysis including techniques such as IDEF (Integrated Computer-Aided Manufacturing Definition Language) and CORE (Controlled Requirements Expression).

(xv) While it is premature to include the models reviewed in system acquisition specifications, they can be used for design evaluation when agreed upon by the procuring agency and the contractor.

(xvi) There is a need to validate many of the existing models and make them operable by a wider variety of users.

(xvii) There is an acknowledged need for models of processes such as planning, decision making, and team performance. This requires additional research to understand these processes.

(xviii) The continuing development of personal computers should encourage the application of many of the existing models.

(xix) Some of the models reviewed have the potential for integration with operational research models, but further work is required to capitalise on this possibility. An extension of such work would be the use of operational research/combat effectiveness models in training and research applications.

(xx) The work of the RSG was not intended to cover models of human response to stress, or the effects of motivation. However, some models were identified which could include (or be modified) to account for such factors, but more work is required on this type of application.

(xxi) Most models require significant user expertise for effective use. The preparation of lecture series and self-tutoring texts is one means to address this problem.

(xxii) Technology demonstrations workshops such as that organised by the RSG are an extremely useful means of technology transfer, but their success is heavily dependent on financial support from organising bodies.

0.3 MAJOR RECOMMENDATIONS

(xxiii) Panel 8 should establish:

(1) an Ad Hoc Group of Experts in 1992 to review progress in modelling and to determine the need for, and structure of, a second technology demonstration workshop;

(2) an Ad Hoc Group of Experts (or Exploratory Group) to prepare recommendations for a NATO Project to implement a library of performance data and models for use in network simulations.

(xxiv) Panel 8 should:

(1) explore with the NATO Military Agency for Standardisation the inclusion of human performance models in weapon system specifications;

(2) encourage collaborative research on cognition, decision making, and team performance;

(3) interact with Panel 7 on the use of human performance models in operational research;

(4) survey member nations for support of a lecture series on model-based methods for analysis, design, and evaluation.

0.4 MILITARY IMPLICATIONS

(xxv) Available training and skill retention models can be used to improve the effectiveness of existing training systems.

(xxvi) Some biomechanical, single task, workload, and multi-operator models can be used to identify critical test and evaluation issues for weapon systems currently in development.

(xxvii) Most models reviewed in this report can contribute to a more cost-effective weapon system design process by facilitating the evaluation of design alternatives. These models allow the designer to incorporate human performance factors that determine weapon system effectiveness.

This page has been left blank intentionally.

OUTLINE OF CHAPTERS

CHAPTER 7 - TRAINING AND SKILL RETENTION MODELS

CHAPTER 8 - NETWORK MODELLING TOOLS

LIST OF FIGURES AND TABLES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND AND PURPOSE

1. In September 1982, the Defence Research Group approved the recommendation of Panel 8 to establish a Research Study Group on "Modelling of Human Operator Behaviour in Weapon Systems". This recommendation was based on the fact that models of human operator performance have been developed with application to a broad range of human behaviours, but have been little used in the design of weapon systems. As stated in the Terms of Reference for RSG.9:

> "Given the potential that such models have for contributing to the analysis, design and evaluation of man-machine systems, there is an obvious need to foster their development and use. An RSG therefore is needed to pull together the available knowledge and stimulate information exchange and co-operative research." (pg. 1)

2. This document is one RSG.9 product addressing this need. It is intended to provide information and recommendations concerning models and model development tools applicable to man-machine system design. It is not a theoretical report and does not address mathematical or conceptual developments required to advance the state of the art. It is meant to be a practical source book for a system designer attempting to locate a model for his or her specific application. In many cases, such a model may not exist. Therefore, the report also reviews model development tools that are readily available and have proven useful in previous applications. No single document can provide the detailed information required to actually apply or develop such models. Therefore, the format is designed to show the designer where and how to obtain that information or expertise. The reference lists will direct the reader to much of the required information. In some cases the report identifies points of contact that may be consulted.

3. By agreement in the Terms of Reference, the report largely limits its scope to models of operator and maintainer performance. There is an emphasis on models that permit some type of computer-based simulation of the man-machine system, as opposed to verbal-analytic or conceptual models. In several cases, it was necessary to broaden the scope. The chapter on training and skill retention, for example, includes several conceptual models which have demonstrated their utility to training system designers. The report also attempts to avoid models which have been used only by their developers in one or two limited applications, and have not demonstrated their design potential.

1.2 GENERAL ISSUES IN HUMAN PERFORMANCE MODELLING

1.2.1 Models: Definitions and Distinctions

4. What are models, and why should system designers be interested in them? At the most basic level, models are nothing more than the specification of functional relationships among sets of variables. The number of variables may range from two to many, and the specification may range from a verbal statement to a precise quantitative formalism. For the

purposes of this report, the functional relations are typically between some system characteristic that the user is interested in manipulating and some aspect of human or system performance that the model user wishes to predict. Considered in this light, the utility of models to the designer is clear. Although creative design involves much more than the knowledge of functional relations, at every point in the design process evaluations of such "if-then" relationships are required.

5. There is disagreement in the modelling literature concerning the need to distinguish between theories and models (Meister, 1985; Pew and Baron, 1983; Sheridan and Ferrell, 1974). While RSG.9 maintains that there are important differences, it is easy to see why the distinction becomes clouded: the most important characteristic of good models and good theories is that they predict behaviour. Nevertheless, the functional relations used to produce these predictions may be quite different in models and theories. Stating that a theory is valid implies more than that the formalism arrives at the right conclusion. In good theory, one is justifiably concerned that the mechanisms expressed in the theory represent the actual processes. Parsimony, and testable hypothesis generation are also important considerations. A useful model, on the other hand, need not be burdened with these requirements. If it predicts behaviour for an area of concern, one may be less concerned with the mechanisms.

6. In spite of this distinction, model developers should be concerned with internal mechanisms. Specifically, we believe that the use of appropriate mechanisms will tend to increase the generality of a model. The use of heuristics will tend to produce a model of limited generality.

7. Although there is disagreement concerning the importance of the model-theory distinction, there is a strong consensus that the appropriate measure of merit for models is utility (Meister, 1985; Pew and Baron, 1982; Sheridan and Ferrell, 1974). That is, do they help to clarify an issue, are they an aid to the user's thinking, do they provide a framework for organising facts, or do they permit the user to experiment with conceptual systems? Meister (1985) states: "The true test of a model is its ability to assist in solving problems, and not necessarily to describe the world in all its details" (pg. 121)

1.2.2 Model Uses

8. Meister (1985) identifies two broad categories of model usage - experimenting with systems and as design aids. He outlines several reasons for the first application:

(1) The system of concern may not be developed to the point that it can be studied directly

(2) Direct study may be too expensive.

(3) The system may be fully occupied or unavailable.

(4) The model may help to simplify phenomena that are too complex to study in the real system.

9. Since the design process involves conceptual experimentation with systems, the above list applies to many design applications as well. For design applications, useful models are nothing more than tools for making tradeoffs at early stages of system development.

10. Although design applications were the focus of RSG.9's work, and of this report, they are not the only use for models. Models are used in the design of experiments, and in the analysis and description of experimental findings. Consider, for example, problems in flight simulation. Given a research question, and the variables to be manipulated, manual control models may be used to select appropriate aircraft dynamics, to select and scale simulated atmospheric turbulence, to select values of the independent variables likely to produce the desired experimental effects, and to select dependent variables which are sensitive to the experimental manipulations. Examples of this type of model usage are given in Levison (1985). Models are useful in the data analysis and description process. For example, the pattern of changes in model parameters, or the differences between model predictions and experimental results can provide insights into the underlying phenomena.

11. A frequently unrecognised application of models is as replacements for humans in operational systems. Automation programmes which replace human control actions, human fault diagnosis, and human decision making all use this technique to some extent. Although the model may not be explicit, the automation scheme will be designed to mimic the strengths of human performance (the model), while attempting to avoid the weaknesses (the non-model aspect of automation). Expert systems applications are an example of replacement models.

12. Models also play a role in education. They can provide simplified accounts, perhaps inadequate for some professional applications, but sufficient to introduce new ideas and relationships to students. Consider Rasmussen's skill/rule/knowledge model (1983) as a technique for broadly categorising behaviour and for suggesting the cognitive strategies associated with each level This model has had a clear impact in education. In this specific example, the model has had significant design usage, as well. The design of large control/display systems, such as those in nuclear power plants, has been affected by Rasmussen's work (1982). Other examples of the educational role of models involve the Crossover Model (McRuer, 1980) and the Optimal Control Model (Baron and Levison, 1980) which have been useful in teaching engineers the concepts of human adaptation to various closed-loop control systems.

1.2.3 General Issues Faced by Model Users

13. The preceding discussion alludes to some of the issues a user must address in selecting or using models First, the user must determine whether the functional relations in the model are applicable to their situation. If the user is considering a model of human lifting behaviour, they must decide whether the lifting geometry described in the model is a reasonable approximation of the lifting technique to be used by their operators. If not, the model predictions may be grossly different from actual lifts. This is the problem of model generality. The model may be highly accurate, but not apply to the user's situation. On the other hand, a model-predicted ability to perform the required lift, with its assumed geometry, may still have significant utility. It may suggest a design modification that will permit the appropriate geometry, a change in operating procedure, or a requirement for teaming on the lift.

14. Implicit in the above paragraph is the issue of model validity. The focus of this issue should be on the ability to predict behaviour in specified situations, not on the "truth" of the assumptions and constructs. Despite this limited definition, few models are adequately tested in this regard. In most cases, what masquerade as problems of validity are actually limitations of generality. The model may be highly accurate in its domain, but the domain is

often poorly defined, particularly in the mind of the user. Pew and Baron (1983) succinctly summarise these points when they argue that the real purpose of model validation is to decide how and under what conditions a model is useful, and whether its usefulness may be improved. The contrast between model generality and validity may seem rather academic to most users. From their perspective, a failure on either count is critical. However, from the perspective of model developers, this is a useful distinction. One would typically invest little effort in modifying an invalid model, but a model of limited generality may be an excellent place to begin such extensions. Limits to model generality are almost always a larger problem than validity. At the simplest level, this is because the model developer has designed his model to be valid in his area of concern. But typically, this area of concern is smaller than that of potential model users.

15. Because of the limited generality of most models, few users will be able to take an existing model and apply it. Although most of the models discussed in this report are well-developed, stand-alone structures, few of them can be directly applied to a specific problem in their published form. In the vast majority of cases some tailoring of the model will be required. Of course, the modelling tools reviewed here, e.g., SAINT, SLAM, HOS will always require specific development for a problem. This fact should not be taken as a criticism of modelling, per se. It simply reflects the state of the art. On the other hand, this fact suggests that users should perform a cost-benefit analysis before investing time in tailoring or extending a model.

16. Interaction effects in human performance are one of the main limitations to model generality. Human behaviour is uniquely affected by combinations of variables, and few models comprehensively include these effects. The severity of this shortcoming depends on the application. For many situations, a model which predicts the main effects of variables will be sufficient. Prediction of main effects may allow the user to identify key variables, and to predict the general effects of design manipulations. In other applications, the interactions are of primary concern. Most large operations research models used in military applications attempt to account for the effects of fatigue, stress, skill level, etc. The difficulty in modelling these particular interactions are often stated as an inability to define performance "modulation" functions. This will continue to be a problem for users. Model developers, although aware of important interactions, often find that they are not well enough defined for modelling purposes.

17. Even in simple cases, users must not expect models to predict the richness of human behaviour. As eloquently stated by Sheridan and Ferrell (1974, pg. 2):

> "People may show grace, imagination, creativity, or feeling even in narrowly constrained tasks; but these qualities are too fine for the nets we cast in modelling and experiment. We have to be content to describe and predict at a more mundane level."

18. In many design applications, prediction at even a mundane level can be quite useful. In some cases behavioural prediction is not required. The designer may simply need assistance in placing the range of human behaviours anticipated with the system into categories that suggest certain control/display approaches. Again, Rasmussen's skill/rule/ knowledge model serves as an example. It proposes particular types of displays for each category of human behaviour (1983).

19. A shortcoming in human performance models is the failure to explicitly model human error. While the reliability models attempt to predict error in a statistical sense, e.g. some number of incorrect switch actuations per 1000 operations, such models are not particularly useful for predicting specific errors with new system designs. This problem has not received a great deal of attention from the modelling community, although there is a cadre of individuals addressing the issue. The interested reader is referred to Chapter 4, Section 4.4 for a good overview of the state of the art in this area.

20. An important area of user concern is the level of expertise needed to use human performance models. Although the requirements vary widely, significant expertise is required for most models. With the decreasing cost and increasing power of personal computers, more and more of these tools are being hosted on small systems. With this trend, RSG.9 has observed significant efforts to make the tools more user friendly. At the present time, required expertise is still a significant limitation to widespread model usage, both in terms of cost and in a lack of understanding of the potential of the tools.

1.3 ORGANISATION OF THE REPORT

21 The remainder of this report is divided into seven chapters. Each of the chapters begins with an overview. The chapter author defines the types of models to be reviewed, summarises the state of the art, and recommends general references known to provide good background material. The author attempts to specify where these techniques are most likely to support good trade-off studies, identify design problems and alternatives, and predict some aspect of ultimate system performance. In so doing, they have attempted to emphasise what these models and tools are likely to achieve in their present forms, not their ultimate potential.

22. The bulk of each chapter consists of model summaries in a format developed by Geer (1976). This format was chosen because of its completeness. It is designed to provide information on the source, purpose, procedures, application examples, limitations, and future needs for each model. Because of format completeness, the user will see some blank areas for several of the models. Although this is unfortunate, it does clearly show what we do and do not know about each of the models reviewed.

1.4 OVERVIEW OF THE REPORT

1.4.1 Chapter 2 - Task Allocation and Workload Prediction

23 This chapter reviews tools which can assist in the allocation of functions to an individual, among the crew members, or to the machine. Since the goal of this process is to make the best use of the abilities of each of these "components", models which suggest optimal allocations would be most useful. However, no such models exist. As a result, this stage is usually an iterative cycle of function allocation .. task analysis .. and workload evaluation to evaluate the designer's proposed solutions. The techniques reviewed in this chapter are more akin to analysis aids than to true models, and are designed to assist the user in the workload evaluation process. Nevertheless, we have included these tools in our report for two reasons:

(1) They typically include implicit models of human workload capacity, serial versus parallel processing, etc.

(2) They constitute the formalised tools available to assist the designer in this process.

24. The chapter reviews only a single model which directly evaluates allocation of functions. The other models found in the literature were either too situation specific or too poorly developed to include. Several models were found for analysing or predicting human operator workload. The most widely used techniques are based on some form of time-line analysis, i.e. comparing the time required to complete the assigned tasks to the time available. The chapter also reviews several recently developed models which address the mental, rather than the temporal, dimension of workload. Although these models are generating a great deal of interest, they have not had much application and validation.

1.4.2 Chapter 3 - Single Task Models

25. This chapter does not pertain to a specific phase of the system design process, although the models reviewed here are most likely to be used in fairly early stages. These models allow behavioural predictions when the human is performing an individual task, e.g. tracking a target, monitoring auditory signals, or making a discrete movement to a control. Because of the simple, constrained task environments, these models contain some of the most formalised mathematical structures reviewed in this report. Many of these techniques can make precise, highly detailed predictions about the performance elements of individual tasks. Often they permit predictions about the effects of molecular equipment characteristics on human performance. They are highly valid in their domain of application. Unfortunately, most of these models have very limited domains. They are not general models of listening, of movement, or of vehicle control.

26. Many of these models have been validated against large sets of data taken in laboratory contexts. Some of the manual control models have been evaluated against real-world data as well. Many of these models have been applied as experiment design and analysis tools, with significant success. However, their use in the design of man-machine systems has been poorly documented, or minimal. The limited documentation of design uses is somewhat understandable. In personal conversations McRuer (1986) described his use of manual control models to assist in the design of flight control systems. In most cases, this work was done under a subcontract to an aircraft manufacturer. It is not surprising that such efforts do not appear in the open literature. The other analytic techniques used in such a process are unlikely to be reported either. Nevertheless, it appears that these models are not commonly used in design. In addition, it appears that when they are used, the work is done by experts, brought in from other companies.

27. Ongoing efforts to host many of these techniques on personal computers, with user-friendly front ends, may improve the utilisation of these tools (see McMillan, et al., 1989., for examples). However, the limited generality of these techniques will continue to be a problem. Several of the models reviewed here can be extended to greater generality. Because these models have been thoroughly tested and their mathematical structures are well understood, they form an excellent basis for extensions by model developers.

1.4.3 Chapter 4 - Multi-Task Models

28. In all realistic environments, humans have multiple task demands competing for their limited perception and action resources. This chapter reviews models which attempt to

predict human performance in some of these complex task environments. The majority are models of display monitoring and decision making in multi-task situations. Some include submodels for rather elaborate control actions based on the decisions, while others only permit selection from a limited set of established control procedures.

29. Most of these models have had minimal validation. This validation has been limited to experimental settings. In these experimental settings, some of these models have made predictions that are useful to designers of supervisory control or command and control systems. However, the generalisation of these results to operational situations has rarely been tested (The principal exceptions are the Human Operator Simulator and the Siegel-Wolf Models). On the other hand, these monitoring and decision models have aided laboratory user's thinking about human strategies in multi-task situations.

30. This chapter reviews few models of human cognitive function or problem solving. This is largely because they are conceptual models, they have not been computerised, and they have tended to remain in the academic and research communities. Another obvious omission is the work in artificial intelligence. Here, NATO has established other groups with this specific charter.

1.4.4 Chapter 5 - Multi-Operator Models

31. The models reviewed in this chapter add a significant level of complexity: communication and interaction among operators. This complexity has allowed these models to be applied to real-world design or evaluation problems. Some of them, such as the Seigel-Wolf model, have been highly successful in a number of applications (Meister, 1985). Most of the models were developed under contracts to military agencies, and were designed to simulate the performance of crews performing specific military missions. Most of these models have an operations research flavour to them. They were not designed to aid in the selection of molecular equipment characteristics, but to address procedural issues, task organisation, task assignments among operators, required crew size, the effects of fatigue or time pressure, etc. In general, these models attempt to predict global system performance measures such as the probability of mission success or the time required to accomplish the mission.

32. Typically, these models require the user to develop rather detailed descriptions of all the tasks to be performed by the crew. This includes descriptions of task criticality, permissible task sequences, average task completion times and variances, and the assignment of tasks to individual crew members. Because of their task analytic nature, operator loading is the key element driving the predictions of several models. Loading is usually defined as the time required to complete all tasks versus the time available. In most cases, these models run in a Monte Carlo, or iterative, fashion to allow sufficient sampling from the task performance distributions to make stable statistical predictions. Success or failure of the mission does not depend on the probability of accomplishing any single task, but on whether or not the crew completes all essential tasks in the required time. Thus, each task has an effect on mission success, but not necessarily a primary one.

33. Some of these models allow the users to evaluate the effects of performance modulation factors such as fatigue and operator skill level. In general, these effects are implemented by modifying the mean and/or variance of task completion time distributions. In come cases, these functions must be provided by the user, while in others such as the Seigel-Wolf model, generalised functions are provided.

34. While the user input requirements are significant, the data often can be obtained at a sufficient level of accuracy from subject matter experts. For complex systems, the resulting models are correspondingly large and complex. These models are rarely, if ever, formally validated by comparing model predictions to the results from actual field trials. Rather, these models tend to be "exercised" and their results evaluated for reasonableness. If the results seem reasonable to the user, design or procedural decisions may be based upon them.

35. These models are often criticised by operational and scientific personnel alike. Such complex models include numerous untested assumptions and approximations. However, it is noteworthy that they continue to be developed and used. Developing and maintaining large operations research models is often a primary mission of studies and analysis offices in military agencies. This suggests that these models satisfy the utility function for managers and decision makers who have few other systematic tools available.

1.4.5 Chapter 6 - Biomechanics and Work Space Design

36. This chapter is also rich in models developed specifically for system design applications, but at a much more molecular level. The review encompasses two principal categories of modelling:

(1) biomechanical models which predict human materials handling capabilities;

(2) anthropometric models which determine the ability of an operator, of a given physical size, to work within a given space, to reach specific controls, and to see specific displays.

37. Biomechanical models which attempt to describe the human body as a mechanical, load-bearing device have a long history. Much of the work in this general area is focused on defining human tolerance limits to vibration and acceleration stress. Such tolerance models are not reviewed in this report. Techniques to model the performance effects of these stressors have also been developed, and some are noted in Chapter 3. In many cases, these performance models are based on existing single-task models with the vibration or acceleration stress represented as a disturbance to visual perception or motor control (Jex and Magdaleno 1978; Korn and Kleinman, 1978).

38. Another type of biomechanical model is the subject of this chapter. These models predict human lifting capacity in materials handling situations. Although this may seem like a restricted domain, reviews of physically demanding tasks in a variety of trades routinely show that most involve lifting (Mital, 1983). In addition, back injuries and back pain continue to be a significant problem for workers performing lifting tasks.

39. Lifting models typically address the issue from one of two perspectives. Some attempt to predict lifting capacity, given specific human, task, and environmental characteristics. Others use Newtonian mechanics to estimate the stresses imposed on the musculoskeletal system during lifting.

40. These models have several significant limitations. Generally, they assume a limited range of lifting postures and geometries, no mechanical aids, smooth symmetrical lifts, good floor contact, and so on. Of course, these assumptions are violated in many operational settings. Nevertheless, these models appear to be receiving increased use.

41. The anthropometric models are computerised versions of traditional drawing board, manikin, and mock-up approaches. As a result, they offer the ability to readily change workstation dimensions and characteristics, to represent individual operator body dimensions, and where appropriate data bases are available to evaluate the fit, reach, and vision envelopes of a wide range of human populations.

42. In many cases, the anthropometric models provide CRT-based graphic output which allows the user to view and interact with the man-models as specified reaches and "looks" are attempted. This feature is an important attribute of the better models as it helps the user to avoid problems such as the man-model reaching through incompletely defined surfaces in the workstation.

43. Although the author of this chapter located references to over a dozen models, most of them are not readily available or widely used. The models reviewed here are widely available and have been used for documented design applications. Of all the models summarised in this report. the anthropometric group is perhaps the nearest to becoming frequently and broadly applied. Despite this fact, they have several common shortcomings. Most have only been validated for a limited range of reaches and fits. Most permit only one or two operator postures, e.g. sitting erect or sitting slumped, and do not represent the effects of postural changes on reach envelopes. Effects of clothing or other restraints are often not modelled. Finally, most models are incompatible with other computer aided design (CAD) programmes and systems.

44. Ongoing developments in anthropometric modelling are addressing many of these limitations. In addition, there are exciting efforts to develop truly dynamic tools that model the actual movement sequences of human operators in proposed workstations.

1.4.6 Chapter 7 - Training and Skill Retention Models

45. This area of human performance modelling has a long and rich history. Much of the model building has been done by psychologists as part of their theory development and testing process. As a result, most of these models are of a qualitative, descriptive nature and were never really intended for system design applications. However, the increasing sophistication of military systems has created a growing need to design better training systems, to forecast their effectiveness during the design process, and to measure their effectiveness during the system life cycle.

46. From the literature on human learning and memory, this chapter has extracted those techniques which have been applied to training issues, or which have been developed to address training system design problems. Although some mathematical models are summarised, several qualitative techniques have been included which have application to the conceptual design of training systems, or to the understanding of human behaviour during the training process. At the current time, these qualitative models have achieved greater utilisation than the quantitative techniques.

47. Perhaps the earliest attempts to describe and predict improved performance with practice involved fitting learning curve data by adjusting the parameters of "fixed-form" equations (Hackett, 1983; Restle and Greeno, 1970). A variety of such equations have been utilised and their relative merits debated. Some success has been achieved in using such models to predict the production rates in industrial settings as workers learned a new manufacturing procedure, for example.

48. The mathematical models developed in the 1940's and 50's by psychological theorists appear to be receiving a second look from model developers. New computer-based numerical methods allow one to more readily identify the parameters of some of these models. Although it is unlikely that these techniques will be used as stand-alone models, they are being used to represent fundamental learning and memory processes in larger models of training and skill retention (Sticha and Knerr, 1983).

49. The qualitative models reviewed in this chapter address a broad range of issues. Some focus on the stages of skill acquisition and have been used in the overall organisation and sequencing of training curricula. Some attempt to categorise human information processing skills and to develop general guidelines for training these skills. Others attempt to predict the effectiveness of proposed training devices on the basis of the difficulty of acquiring certain skills and the importance of these skills to the operational tasks. These and other qualitative models are being used in the training system design process, and are a significant area of model development activity.

1.4.7 Chapter 8 - Network Modelling Tools

50. This chapter reviews special purpose computer languages developed for the purpose of simulating man-machine systems. There is an important tie between the tools reviewed here and the models reviewed in Chapter 5 - Multi-Operator Models. Many of those multi-operator models have been developed using simulation languages such as SAINT, SLAM, or Micro-SAINT. As noted by the author of this chapter, this is not by chance. Our theoretical foundations are weak in the area of group-interactive behaviour. As a result, it is difficult to develop formal mathematical models of crew performance. In general, the theoretical constructs are not available for developing equations that describe the performance of multi-operator systems.

51. On the other hand, such systems can be simulated as task networks which represent the sequence of activities in the system. In setting up such a network, the user must specify for each task: (1) the predecessor tasks that must be completed before the task in question can begin, (2) the statistical characteristics of the task, and (3) the branching to other tasks to be performed upon task completion. The statistical characteristics of a task may include task duration distributions, task criticality, operator speed, operator accuracy, etc. Different operators may be responsible for different tasks, and probabilistic branching among tasks may be used to represent interactions among operators.

52. These modelling tools provide a flexible structure for simulating the performance of large-scale systems. With this flexibility comes the requirement to provide a large body of data about the system's structure and task characteristics. In many cases, such data can be obtained by interviewing subject matter experts. As the chapter author notes, validation of such large-scale models is almost impossible, and any specific model is not likely to generalise to other systems. However, the same constraints become even more severe when one attempts to

evaluate such systems experimentally. Such research is extremely difficult.

53. While these models can be criticised on the grounds of validity and generality, they may provide the user's only option for systematically evaluating alternative system designs. Although the use of these tools is constrained by our limited ability to specify the statistical characteristics of tasks, to define the interactions among tasks, and to decompose a complex system into tasks, what better option is available to the user? In the opinion of RSG.9, the simulation models which can be developed with these tools have far greater utility for addressing the design of complex, multi-person systems than any of the formal, mathematical models reviewed in the previous chapters.

54. Because of the utility of these techniques, the current emphasis on hosting them on inexpensive workstations and personal computers is very encouraging. There also appears to be a parallel emphasis on making these techniques easy to learn and use.

1.5 A BRIEF LOOK AT THE FUTURE

55. The trends noted for the modelling tools are most important. As long as human performance models are difficult or expensive to access, and considered the domain of highly-trained experts, they will not be used. Consider some of the activities now routinely accomplished on personal computers. Spreadsheets, complex database programmes, and desktop publishing are now performed by people with little training or special expertise. Only a few years ago, these activities were limited to specialists. Powerful, user-friendly software has made these tools highly successful and usable. Small dollar problems can be addressed with these tools in a cost-effective manner. Clearly, model developers must follow the lead of successful software developers if they want their models to be broadly used as design aids.

56. Fortunately, this is happening. Early in our life cycle, RSG.9 considered sponsoring the development of some models on micro-computers. However, a brief survey of the field showed that it was already taking place. A significant number of developers have produced, and are continuing to develop, commercially available models with serious attention to cost reduction and ease of use. These developments are specifically reviewed in each of the report chapters.

57. Given that these trends will make models more accessible and usable, which techniques are likely to enter the designer's "tool-kit" in the future? First, it seems certain that the anthropometric models will continue to be improved and integrated into CAD/CAM systems. Second, the use of the network simulation languages will continue to expand. Their cost is becoming very reasonable, and they are part of the curriculum in many universities. The current emphasis on workload reduction in military systems would suggest that models for workload analysis will also become readily available. However, none of the current time-line analysis tools seem destined to play this role. Perhaps the simulation languages will be utilised to develop individual models for specific purposes. The training models will continue to play an important role in the development and analysis of training systems.

58. But what of the many highly mathematical models of individual tasks, or of performance in multi-task environments? It does not appear that these models will have a significant near-term impact on design. Such models will continue to be developed and evaluated in universities and laboratories engaged in theoretical pursuits. There is ongoing work to host some of these models on microcomputers with improved user interfaces. While

this will greatly improve their accessibility, it appears that the primary users of these models will be experts. It is likely that these models, or their basic concepts, will be incorporated in larger models developed using the simulation languages. Here, these models can fill a significant gap in system simulations, which desperately need descriptions of task and subtask performance.

59. Does this prognosis paint a bright picture for the application of human performance modelling to design? RSG.9 believes that it does. It signals a shift in thinking about models as relatively discrete, off-the-shelf packages to thinking about modelling software that allows the user to access, modify, and use the above components to construct simulations at a level of complexity appropriate for their design problem. Obviously, we have some distance to go in achieving this objective. But hopefully, improved user interfaces, simulation languages for description of the overall system structure, and task component models to fill in the details, will together provide the necessary tools.

1.6 REFERENCES

Baron, S., & Levison, W. H. (1980). The optimal control model: status and future directions. Proceedings of IEEE Conference on Cybernetics and Society. New York: Inst. of Electrical and Electronic Engineers.

Geer, C. W. (1976). Navy managers guide for the analysis sections of MIL-H-46855 Boeing Aerospace Company (Report D 180-19476-2) Warminster, PA: US Naval Development Center.

Hackett, E. A. (1983). Application of a set of learning curve models to repetitive tasks. The Radio and Electronic Engineer, 53, 25-32.

Jex, H. R., & Magdaleno, R. E. (1978). Biomechanical models for vibration feed-through to hands and feet for a semi-supine pilot. Aviation, Space, and Environmental Medicine, 49, 304-317.

Korn, J., & Kleinman, D. L. (1978). Modelling the effects of high-g stress on pilots in a tracking task. Proceedings of the Fourteenth Annual Conference on Manual Control (NASA Conference Publication 2060, pp. 55-62). Moffett Field CA: National Aeronautics and Space Administration/Ames Research Centre.

Levison, W. H. (1985, October). Application of the optimal control model to the design of flight simulation experiments (Paper 851903). SAE Aerospace Technology Conference Proceedings, SP-634 - Flight Simulation/ Simulators. Warrendale, PA: Society for Automotive Engineering.

McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

McRuer, D. T. (1980). Human dynamics in man-machine systems. Automatica, 16, 237-253.

McRuer, D. T. (1986). Personal communication.

Meister, D. (1985). Behavioural analysis and measurement methods. New York: Wiley.

Mital, A. (1983). Preface to the special issue on manual materials handling. Human Factors, 25, 471-472.

Pew, R. W., & Baron, S. (1983). Perspectives on human performance modelling. Automatica, 19, 663-676.

Rasmussen, J. (1982). The role of cognitive models of operators in the design, operation, and licensing of nuclear power plants. Proceedings of the Workshop on Cognitive Modelling of Nuclear Plant Control Room Operators (NUREG/CR-3114, ORNL/TM-8614). Oak Ridge, TN: Oak Ridge National Laboratory.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 257-266.

Restle, F. & Greeno, J. G. (1970). Introduction to mathematical psychology. New York: Addison-Wesley.

Sheridan, T. B., & Ferrell, W. R. (1974). Man-machine systems: Information, control, and decision models of human performance. Cambridge, MA: The MIT Press.

Sticha, P. J., & Knerr, C. M. (1983). Task-element and individual differences in procedural learning and retention: A model-based analysis (Report DDI/PR 83-1-334). McLean, VA: Decisions and Designs Inc.

This page has been left blank intentionally.

CHAPTER 2

TASK ALLOCATION AND WORKLOAD PREDICTION

2.1 INTRODUCTION

60. Human Factors/Ergonomics approaches to the design of systems emphasise the importance of systematically allocating functions to man or machine. That process, formalised as 'Function Allocation Analysis', ultimately determines the tasks that will be performed by the human operators and maintainers of the system (see Geer, 1976; Meister, 1985; Parks, 1987 for example).

61. There are few formal techniques for conducting such analyses. Geer (1976) identifies three ways of performing a Function Allocation Analysis:

(1) trial and error substitution of each function allocation option into a system or subsystem model;

(2) an evaluation matrix technique based on qualitative (ordinal) performance data such as the Fitts' List;

(3) a candidate design evaluation matrix, in which sub-system functions are listed and candidate function allocations compared on the basis of performance criteria such as response time, error rate, operability, cost, etc.

62. Models of human performance could contribute to comparisons of performance criteria, if they could be related directly to system performance. This is not always possible, however, and system effectiveness is often addressed through the intervening variable of 'operator workload'. This approach is based on the premise that the human operators have a finite performance capacity, which, if exceeded, will lead to performance degradation. The approach uses an iterative cycle of:

Function Allocation - Task Analysis - Workload Analysis - Function Re-allocation

to ensure that the human components of a system are not overloaded, and are performing tasks which make best use of their abilities.

63. Workload models are being used increasingly, because they permit the exploration of systems concepts early in the system development process, before man-in-the-loop simulation facilities are available. Some users report quite high correlations between workload model predictions and operator performance, or operator subjective workload estimates. Parks & Boucek (1989) report greater than 90% accuracy in the performance prediction of a time-line model and actual aircraft flight data. Holley (1989) also reports time-line predictions which were within 10% of times in a flight simulator or actual flight. Correlations between the workloads predicted by such models and subjective workload estimates tend to be lower. Parks & Boucek (1989) report a correlation of +.63 between subjective workload estimates and the model prediction. Bateman & Thomson (1989) reported an average correlation of +.74

between subjective workload estimates in a simulator and the predictions of a time and intensity model. Potential users should note that there are circumstances in which operator subjective workload and task performance can disassociate (Hart & Wickens, 1990).

2.2 OVERVIEW AND RECOMMENDED REFERENCES

64. Some of the models reviewed in other sections of this report could contribute to Function Allocation Analysis, but few appear to have been used for such purposes. Only one model has been identified which is intended specifically for such use. The Function Allocation Model (FAM, Section 2.3.1) is based on the candidate design evaluation matrix technique outlined above (Geer, 1976). FAM approaches function allocation by evaluating a proposed configuration using the metric of operator workload.

65. Workload Analysis is directed to validating the function allocation in terms of the resultant load on the operator. Workload is a multi-faceted concept (see, for example, Gopher & Donchin, 1986; Hancock & Meshkati, 1988; Hart & Wickens,1990), and a variety of approaches have been taken to workload modelling (Linton, Plamonden, Dick, Bittner, & Christ, 1989). At the least, workload is a function of the physical, temporal, and mental demands placed on the operator (Hart and Staveland, 1988). The approaches to operator workload reviewed in this chapter emphasise the temporal and mental demands of the operator's tasks, rather than the physical. Some models of physical workload are reviewed in Chapter 6.

66. In general, the most widely used model for workload analysis appears to have been the time-line (temporal demand) based approach. The simplest approach calculates workload on the basis of:

time required for tasks / time available x 100

for sequential 'mission time segments' (time intervals of a few seconds or a minute). This model is related to models for studying the manning of multi-operator systems (Moore, 1971) and Monte-Carlo simulations of manned systems (Lozano, Albanese & Sherwood, 1978; Plato, 1974; Underwood & Buell 1975, for example).

67. At its most basic level, it is debatable whether time-line analysis incorporates a model of human performance, in the sense of a formal description or abstract representation. The only 'model' is that implicit in the concept of operator capacity being related to time 'stress'. This concept was extended by relating the probability of successful performance to the time-line workload (Jones & Wingert, 1969). Belcher (1973) reported a different approach, in which ratings of workload for individual tasks, produced independently, were used to determine whether the operator was overloaded and subsequent tasks should be delayed. The Siegel-Wolf models discussed in other sections of this report use a more elaborate version of that concept, relating the probability of correct performance of a task to the time stress created by a sequence of tasks. A further elaboration of the time-line model of workload was the development of the 'functional interlace' concept, which allowed for the partial overlapping of some tasks (Wingert, 1973).

68. The increasing trend towards employing human operators of new systems in monitoring, supervisory, and executive tasks emphasises the mental demand aspects of workload. A variety of modelling approaches have been developed which address mental

demand. Some of the concepts are similar to time-line analysis techniques. For example, the approach to decision making workload developed by Tulga and Sheridan (1980) uses a 'loading factor' of 'time required to do an average task/ time affordable to do an average task'. Other developments relate workload to the amount of simultaneous activity, or attentional demand, e.g. eye only, eyes and hand, hand only, etc. Such partitioning of workload has also been used in some time-line models (see Time Line Analysis and Prediction 2.3.2; WAM 2.3.3).

69. In the 'attentional demand' approach, workload is determined from the sum of each class of activity on the mission time-line, (Aldrich, Szabo, & Bierbaum, 1989; Meister, 1985). This approach has been elaborated to avoid the simple addition of attentional demands, through the use of rules for their combination (see Attentional Demand Model 2.3.6; W/INDEX 2.3.7). The latter model includes task conflict 'cost' functions which are, in a sense, the reciprocal of the 'functional interlace' allowances developed by Wingert (1973) to deal with time sharing of tasks.

70. The models reviewed in this chapter differ in their approach to the concept of workload, to the variables which they include in their calculation of workload, and in the task details which they incorporate. Some models can be used early in concept development, when operator tasks are known, but not defined. Other workload models require details of the human-machine interface. Potential users should review any proposed application with care, to ensure that they will have the information necessary to run the model at the time they need it in the system development process.

71. One class of workload models not covered in this review is that related to 'dynamic re-allocation of function'. Continuously adapting the allocation of functions to maintain an 'optimal' operator workload has been studied intermittently since the late 1960s. The increasing capabilities of hardware and software, and the resultant shift in operator's functions to monitoring and supervising, encourage such an approach. There have been several attempts at studying 'dynamic re-allocation of function', using models of system performance. No one model has emerged as the most promising approach for such a technique, however. For example, Rouse (1980) reports modelling approaches to adaptive function allocation based on queuing theory. The models assume a system with two servers, one automatic, one human, and a class of task 'customers'. The models appear to be highly situation specific, and they are probably better suited to subsequent stages of analysis/design.

2.2.1 References

Aldrich, T.B., Szabo, S.M., & Bierbaum, C.R. (1989). The development and application of models to predict operator workload during system design. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, and L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Bateman, R.P., Thomson, M.W. (1989) Correlation of predicted workload with actual workload measured using the subjective workload assessment technique. Poster presentation to NATO Workshop on Applications of Human Performance Models to System Design. Wichita, KS: Boeing Military Airplane Co.

Belcher, J.J. (1973). A technique for assessing operability/effectiveness of control-display systems. In K.D. Cross and J.J. McGrath (Eds.), Crew System Design: An Interagency Conference. Santa Barbara, CA: Anacapa Sciences Inc.

Brown, E.L., Stone, G., & Pearce W.E., (1975). Improving cockpits through flight workload measurement. Proceedings of the 2nd Advanced Aircrew Display Symposium. (Douglas paper 6355). Patuxent River, MD: US Naval Air Test Center.

Geer, C.W. (1976). Navy manager's guide for the analysis sections of MIL-H-46855. (Boeing Aerospace Co. Report D180-19476-2), Warminster, PA: US Naval Air Development Center.

Gopher, D, & Donchin, E. (1986). Workload - an examination of the concept. In K.R. Boff, L. Kaufman, J.P. Thomas (Eds.) Handbook of perception and human performance. New York: John Wiley & Sons. Chapter 41.

Hancock, P.A., & Meshkati, N. (Eds.) (1988). Human mental workload. Advances in Psychology, 52. Amsterdam: North Holland B.V.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.) Human mental workload. Advances in Psychology 52. Amsterdam: North Holland B.V.

Hart, S.G., Wickens, C.D (1990). Workload assessment and prediction. In H.R. Boohrer (Ed.) MANPRINT: An approach to systems integration. New York: Van Nostrand Reinhold.

Holley, C.D. (1989). A model for performing system performance analysis in predesign. G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Jahns, D.W. (1972). Operator workload: what is it and how should it be measured? In K.D. Cross and J.J. McGrath (Eds.), Crew System Design: An Interagency Conference. Santa Barbara, CA: Anacapa Sciences Inc.

Jones, A.L., & Wingert, J.W., (1969). On estimating the capability of an avionic man-machine system. International Symposium on Man-Machine Systems. (IEEE Conference Record No. 69C58-MMS).

Linton, P.M., Plamondon, B.D., Dick, A.O., Bittner, A.C., & Christ, R.E. (1989). Operator workload for military system acquisition. In G.R. McMillan, D.Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Lozano, P.A., Albanese, R.A., & Sherwood, S., (1978). Crew manning in long duration airlift and tactical settings. Proceedings, ASCC Working Party 61 Symposium on Aeromedical Factors in Air Combat Capabilities. Canberra, Australia.Air Standardization Coordinating Committee.

Meister, D. (1985). Behavioral analysis and measurement methods. New York: John Wiley & Sons.

Moore, C.B. (1971). Space mission modelling and simulation techniques. Biotechnology - Proceedings of a Conference. (NASA SP-205) Washington D.C.: NASA

Parks, D. (1978). Current workload methods and emerging challenges. In N. Moray (Ed.), Mental workload - its theory and measurement. New York: Plenum Press.

Parks, D. & Boucek G.P. (1989). Workload prediction, diagnosis, and continuing challenges. G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Parks, R.E. (1987). Development and testing of a predictive methodology for optimization of man-machine interface in future avionics systems. The Design, Development, and Testing of Complex Avionics Systems. (AGARD-CP-417), Neuilly-sur-Seine, France: Advisory Group for Aerospace Research and Development.

Plato, A.I., (1974). The naval engineer's role in manpower requirement development for new ships. Proceedings, 11th Annual ASE Symposium. Washington, D.C.

Rouse, W.B. (1980). Systems engineering models of human-machine interaction. New York: North Holland.

Tulga, M.K., & Sheridan, T.B., (1980). Dynamic decisions and work load in multitask supervisory control. IEEE Transactions on Systems, Man, and Cybernetics, SMC-10 (5), 217-232.

Underwood, Cdr. F.S., & Buell, G.D., (1975). Computer-aided pilot house design - a systems approach. Proceedings, Fourth Ship Control Systems Symposium. den Helder, Netherlands: RNIN.

Wingert, J.W., (1973). Function interlace modifications to analytic workload prediction. In K.D. Cross, & J.J. McGrath (Eds.). Crew System Design: An Interagency Conference. Santa Barbara, CA: Anacapa Sciences Inc.

2.3 MODEL SUMMARIES

2.3.1 Function Allocation Model (FAM)

Summary Description

72. FAM is a collection of computerised algorithms which will process and test different options for allocating functions to operators or equipment against a given set of mission requirements. FAM works from lists of performance functions, performance data and allocation options to evaluate promising allocation options.

73. FAM contains two major data processing routines:

(1) Mission Evaluator, which computes the probability of overall mission success for various function allocation candidates, and for specific mission objectives if required;

(2) Procedure Generator, which derives data on operational procedures and procedure statistics; those data can be used to document and refine operational sequence diagrams.

History and Source

74. The model is part of the Computer Aided Function Allocation and Evaluation System (CAFES) package developed for the US Navy by Boeing Aerospace Company (see References). The impetus for development of CAFES was a requirement to standardise on the analytic processes used in human engineering analyses (as dictated by US MIL-H-46855, for example), and to provide a conceptual framework for such analyses which would lead systematically from one analytical step to the next, and would be amenable to the use of computer aids. CAFES is thus a complete analytical package, of which FAM is a module. All CAFES modules were intended to be used either individually or in combination with other modules. The CAFES software was developed at Boeing, and transferred to computing facilities at US Naval Air Development Center (Naval Air Development Center).

Product and Purpose

75. The overall objective of FAM is to identify and rank order function allocation schemes for systems by performance effectiveness.

76. The Mission Evaluator module produces:

(1) estimates of the reliability of achieving mission objectives,

(2) estimates of overall mission reliability,

(3) estimates of perceived task load,

(4) integrated task reliabilities.

77. The Procedures Generator module produces estimates of:

(1) percent of tasks interrupted and completed,

(2) percent of tasks interrupted and not completed,

(3) percent of tasks not started,

(4) percent of tasks started but not completed,

(5) percent of tasks started late,

(6) percent of time operator busy,

(7) tasks interrupted and completed,

(8) tasks not started,

(9) task simultaneity status,

(10) tasks started late.

78. In addition FAM is designed to produce a rank order of candidate function allocation combinations and data for the production of operational sequence diagrams.

79. FAM is intended for the study of:

(1) optimum task allocation,

(2) most effective level of automation,

(3) best crew size.

When Used

80. It is claimed that FAM, as part of CAFES, can be used iteratively throughout concept development. The amount of detail required as input to the model suggests that it could not be used from the outset of a completely new development.

Procedures for Use

81. The user must have access to CAFES, and must prepare the following data for the programme:

(1) detailed operational requirements,

(2) mission scenarios,

(3) function flow block diagrams,

(4) mission time-lines,

(5) task analyses,

(6) task performance data,

(7) task allocation concepts.

82. Note that the task performance data would reflect assumptions about user training level and skill.

Advantages

83. Manual methods of function allocation are so laborious that they discourage the development and evaluation of more than one design concept. FAM (in fact the whole CAFES package) is intended to encourage the evaluation of different concepts by reducing the manual work required.

84. In addition, as with other decision aids, it is claimed that FAM permits the factors having most effect on system performance to be identified, and the assumptions which lead to the model outcome to be critically examined.

Limitations

85. FAM requires input of a great deal of data. Seven different data classes are specified, each containing several types of detailed data. Examples are:

86. Mission Scenario Data:

(1) tasks,

(2) number of task occurrences,

(3) start time and duration for each occurrence.

87. Task Performance Data:

(1) operator reliability as a function of task execution time,

(2) nominal task execution time,

(3) equipment reliability,

(4) task load ratings,

(5) task priority,

(6) task interruptability classification.

88. It was intended that much of these data would be made available by preceding analyses carried out by other CAFES modules. It was also intended that the data input load would reduce as successive projects were undertaken, because of redundancy and commonality between successive systems. One of the assumptions of the overall CAFES development was that a full set of representative function flow analyses for a variety of systems (aircraft, ships, command and control) would be built up. Therefore, it would be more difficult to use FAM either independently of CAFES, or for a series of different projects. Some users report that CAFES cannot be transported away from the computer on which it was developed, thereby restricting its availability.

Application Examples

89. No information on applications of FAM has been obtained at the time of preparation of this review. Some users have suggested that not all modules of CAFES have been run satisfactorily.

Technical Details

90. No information on the technical details of FAM has been obtained.

References

Curnow, C.W. & Ostrand, R.A. (1978). Function allocation model, FAM programmer manual. (Report D180-20247-6). Seattle, WA: Boeing Aerospace Company.

Geer, C.W. (1981). Human engineering procedures guide. (AFAMRL-TR-81-35). Dayton, Ohio: U.S. Air Force Aerospace Medical Research Laboratory.

Hutchins, C.W., (1974). A computer aided function allocation and evaluation system (CAFES). Proceedings of the 18th Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society.

Parks, D.L., & Springer, W.E., (1975). Human factors analytic process definition and criterion development for CAFES (Report D180-18759-1). Seattle, WA: Boeing Aerospace Company.

Future Needs

91. According to Parks and Springer (1975) the model can be further developed in terms of the performance data which it handles. "The task information currently used is derived from 'expert' knowledge, including knowledge of data sources, rather than from a storage bank of information. The ability to further automate task allocation and workload analysis requires the inclusion of stored information about human movement, operation, manipulation, cognition, processing and reactive times as well as efficiency, and effectiveness or reliability of the above definitions."

92. Information from one source responsible for the FAM development suggests that it requires improvements in its ease of use. Given its development history, it appears likely that the overall CAFES development may have been superseded by the development of other, commercially available, software.

2.3.2 Time-line Analysis and Prediction (TLAP)

Summary Description

93. The time-line approach to workload description was reviewed in the introduction to this chapter. The model addresses the temporal demands of workload, through the concept of time stress. Thus workload is expressed as the ratio:

Tr / Ta x 100

where Tr is the time required to perform a sequence of tasks, and Ta is the time available. Usually the model includes a concept of time-stress in the form of a capacity limit of 70 to 80% occupied (Jones & Wingert, 1969; US Department of Defense, 1987). Thus if the calculated workload is 85%, it is regarded as marginal and it can be expected that, in practice, the operator would shed some of the less important tasks to retain a level closer to 80% (see Parks, 1978; Parks & Boucek, 1989).

94. Parks and Boucek (1989) describe an elaboration of the model which deals with mental demands on the operator. This approach uses a complexity score for estimating cognitive workload, based on information theory. The 'information content' for each display and control is defined as the number of 'bits' (binary digits) into which they are encoded, depending on the number of alternatives they present.

History and Source

95. Time-line analyses are related to approaches to workload balancing used in Methods Study and Industrial Engineering. The model has been in widespread use for a number of years, in both North America and Europe. Use of the model has been described by Jones and Wingert of Honeywell Inc. (1969), Jahns (1972), Brown, Stone and Pearce (1975) of Douglas Aircraft Inc., Chiles (1978), Parks (1978) of Boeing, who traced that company's use of the model back to 1959, and US Department of Defense (1987).

Product and Purpose

96. The model produces an estimate of operator workload against the mission segment time line. As Parks and Boucek (1989) describe, additional information can be added to this basic output, including a 'red line' for unacceptably high levels, and lists of tasks and sub-systems associated with excessive workload.

When Used

97. The model is intended for use early in system concept development, once a detailed task analysis has been completed. It forms part of the cycle of

function allocation - task analysis - workload prediction - function re-allocation

mentioned in the introduction to this chapter. The extension of the model to include cognitive tasks (Parks & Boucek, 1989) requires a detailed description of displays and controls. This would preclude its use in the earliest stages of concept development, unless an evolutionary approach to design was being followed from one system to the next.

Procedure for Use

98. The user must develop detailed task analyses, indicating the start and stop times, or the start time and duration of each task, as well as their sequences. Elaborations of the model require additional data. The length of the time 'window' for the Tr/Ta calculation must be specified. The workload is calculated on a window by window basis, as indicated in the summary description.

Advantages

99. Jahns (1972) has described time-line approaches to workload estimation as useful techniques for making broad predictions about an operator's ability to perform a given set of tasks. Time-line analysis has the advantage that it can be related directly to the system missions and mission analyses. It is therefore highly suited to applications where time stress is likely to arise due to external pacing events.

100. As described here the model is adaptable, and can be used early in concept definition, and refined and expanded as the design becomes more detailed.

Limitations

101. The model requires a comprehensive bank of time data for all the operator tasks. Both Wingert (1973) and Chiles (1978) have reported that the time-line approach can overestimate the time required for a sequence of tasks. Those reports are at variance with those of Parks (1978) and Parks and Boucek (1989). The model appears to require a high level of skill in its application, particularly in the treatment of 'continuous' tasks. For the novice user, the problem of representing tasks which are performed concurrently with continuous tasks is not easy to deal with. Parks and Boucek suggest that concurrent tasks are best dealt with by splitting the model into distinct channels (internal vision, external vision, left hand, right hand, left foot, right foot, hearing, speech, and cognition). Such partitioning is described in WAM, 2.3.3, and subsequent models.

102. As Jahns (1972) noted, time-line models are highly deterministic. That is, they assume that the operator(s) will perform tasks in one set sequence, and they seldom include any variance in operator performance or mission sequences or events.

Application Examples

103. Such models have been widely used, particularly in the aerospace industry (see Geer, 1976; Jones & Wingert, 1969; Parks, 1978; Parks & Boucek, 1989). Parks and Boucek report validation studies in which predictions made from the model were compared with data taken from the Boeing 757 and 767 aeroplane projects. They report greater than 90% agreement. They also report the results of a joint Boeing/McDonnell Douglas Aircraft Company study, in which time-line workload predictions were compared with 12 other workload indices.

Technical Details

104. The model is included in some available software packages, but can be programmed easily on most computers. Some large computer analyses have been run using this model (see Brown, Stone, & Pearce, 1975, for example).

105. The time period (window) for which the Tr/Ta ratio is calculated is obviously important. In the one extreme, an instantaneous ratio will calculate either 0 or 100% workload: in the other extreme, the ratio for a complete mission segment will average out the unacceptably high workload peaks which the analysis is intended to identify. General practice appears to be to calculate Tr/Ta using a moving window of seven to ten seconds. Additional technical developments are described by Parks and Boucek (1989).

References

Brown, E.L., Stone, G. & Pearce, W.E. (1975). Improving cockpits through flight workload measurement. Proceedings of the 2nd Advanced Aircrew Display Symposium. (Douglas Paper 6355). Patuxent River, MD: US Naval Air Test Center.

Chiles, W.D. (1978). Objective methods. Assessing Pilot Workload. (AGARD-AG-233, Chapter 4), Neuilly-sur-Seine, France: Advisory Group for Aerospace Research and Development.

Geer, C.W. (1976). Navy manager's guide for the analysis sections of MIL-H-46855. (Boeing Aerospace Co. Report D180-19476-2), Warminster, PA: US Naval Air Development Center.

Jahns, D.W. (1972). Operator workload: what is it and how should it be measured? In K.D. Cross & J.J. McGrath (Eds.). Crew System Design: An Interagency Conference. Santa Barbara, CA: Anacapa Sciences Inc.

Jones, A.L. & Wingert, J.W. (1969). On estimating the capability of an avionic man-machine system. Proceedings, International Symposium on Man- Machine Systems. (IEEE Conference Record No. 69C58-MMS).

US Department of Defense (1987). Military handbook: human engineering procedures guide (DOD-HDBK-763), Washington, D.C: US DoD.

Parks, D. (1978). Current workload methods and emerging challenges. In N. Moray (Ed.), Mental workload - its theory and measurement. New York: Plenum Press.

Parks, D. & Boucek G.P. (1989). Workload prediction, diagnosis, and continuing challenges. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Wingert, J.W., (1973). Function interlace modifications to analytic workload prediction. In K.D. Cross Jr, & J.J. McGrath (Eds.). Crew System Design: An Interagency Conference. Santa Barbara, CA: Anacapa Sciences.

Future Needs

106. Parks and Boucek (1989) argue for continued evolution and refinement of the model. They suggest several possible developments which could be implemented. In addition, the deterministic approach to such models needs to be replaced by models which reflect the variance in operator task times and task sequences that are observed in actual operations. The development of the SAINT modelling language (see Chapter 8) provides the possibility of representing probabilistic mission segments and events, and varying operator task times.

2.3.3 Workload Assessment Model (WAM)

Summary Description

107. This model is part of the Computer Aided Function Allocation and Evaluation System (CAFES) programme (see FAM, 2.3.1). It is a specific application of a time-line (TLAP type) workload model, based on the ratio of the time required to time available to perform sequential mission tasks. The model treats the human operator as a set of channels (eyes, hands, feet etc.). WAM requires inputs of data for task sequences, nominal task times, task time per channel, and details of operator and equipment allocations. The model calculates time per channel per task event, percent of time each channel is busy during each time segment, means and standard deviations of workload for mission segments; it identifies those periods when the workload exceeds a 'critical workload threshold', and tabulates and plots the above information.

History and Source

108. As a module of CAFES, WAM was developed for US Naval Air Development Center by Boeing Aerospace Company. One of the earliest references to WAM was in the context of ship system development (Whitmore, 1975). As outlined in the review of the Function Allocation Module (FAM) model (2.3.1), the impetus for CAFES was the need to standardise on the analytic processes used in human engineering analyses which would be amenable to computer aiding. The software was developed at Boeing and transferred to Naval Air Development Center.

Product and Purpose

109. WAM produces tabulations and plots of operator workload for mission tasks and mission segments, broken down into:

(1) external vision,

(2) internal vision,

(3) right hand,

(4) left hand,

(5) right foot,

(6) left foot,

(7) cognitive,

(8) auditory,

(9) verbal.

110. The results can be compiled for activities associated with a given operator, given sub-system or given mission segment.

When Used

111. WAM is intended to be used for evaluating operator workload to validate previous Function Allocation Analyses. As such it can be used iteratively throughout system development.

Procedures for Use

112. WAM requires data on mission phases and task sequences, including times for each phase, based either on analysis, experience or experimentation. The mission phases are then divided into a time-line analysis of small time intervals (six second intervals are recommended), with all tasks identified and named. Estimates of 'channel utilisation time' must then be prepared showing the time in seconds that each channel is used during each time-line segment. The data are then input to the programme and the output formats selected.

Advantages

113. The developers claim that WAM produces a more objective review of operator workload than previous all-manual methods. The various formats can be used to highlight workload associated with specific items of equipment, specific tasks or operators, or to highlight those activities when workload exceeds a given critical level.

Limitations

114. WAM is entirely dependent on the input data, particularly on the user's estimates of operator channel workload. Those data are time-consuming to prepare, and, obviously, the programme is most effective when used in conjunction with the other CAFES modules which include some of the required data. The problem of representing a mixture of continuous and intermittent tasks mentioned in the review of Time Line Analysis and Prediction (2.3.2) appears to have been dealt with by ignoring the continuous tasks (see Linton, Jahns, & Chatelier, 1977). This simplifies the construction of the model, but brings into question the extent to which time-sharing contributes to operator workload.

Application Examples

115. Linton, Jahns, and Chatelier (1977) describe an application of WAM to the prediction of pilot workload in two missions - deck-launched intercept and close air support. The paper states that the WAM model (referred to as SWAM) had been validated previously by Boeing for both military and civil applications. No other references to applications were found.

Technical Details

116. No technical details of the WAM model have been obtained at the time of writing.

References

Curnow, R.P., Edwards, R.E., & Ostrand, R.A., (1977). Workload assessment model, WAM programmer manual. (Report D180-20247-2), Seattle, WA: Boeing Aerospace Co.

Edwards, R.E., Curnow, R.P., & Ostrand, R.A., (1977). Workload assessment model, WAM user manual. (Report D180-20247-3), Seattle, WA: Boeing Aerospace Co.

Linton, P.M., Jahns, Dieter W., & Chatelier, Cdr. P.R., (1977). Operator workload assessment model: an evaluation of a VF/VA-V/STOL system. Methods to Assess Workload. (AGARD-CP-216, paper A12), Neuilly-sur-Seine, France: Advisory Group for Aerospace Research and Development.

Parks, D.L., & Springer, W.E. (1975). Human factors engineering analytic process definition and criteria development for CAFES. (Report D180-18750-1). Seattle, WA: Boeing Aerospace Company.

Whitmore, D.C., (1975). CAFES applications in ship system development. Seattle, WA: Boeing Aerospace Co., Crew Systems Div.

Future Needs

117. The programme would obviously benefit from the inclusion of a task time data base, which is implied in the developments recommended for CAFES by Parks and Springer (1975). It appears likely that this development will be superseded by other software packages currently being developed, which will be more accessible and easier to use.

2.3.4 Functional Interlace Model

Summary Description

118. This model uses a time sharing approach to the time required vs time available metric of operator workload. The model is based on the postulation that sensory inputs and motor outputs are interlaced through parallel processing. The extent of parallel processing is dependent on the extent of 'incompatibilities' in the operator's sub-tasks. The model is in many senses a precursor to more recent developments based on the concept of the human operator having multiple resources which can be drawn upon without mutual interference.

History and Source

119. The model was developed by Wingert at Honeywell Inc. It appears to have been developed to meet the limitations of other time-line workload analyses previously used at the company (Jones & Wingert, 1969), which were similar to TLAP(2.3.2) and WAM (2.3.3). Wingert (1973) reported that simple time-line models were found to over-estimate the time required to complete a sequence of sub-tasks. Others have reported similar problems of over over-estimation of workload in time-shared tasks - see Chiles, 1978 for example. The Functional Interlace model was reported at the US Interagency Conference on Crew Station Design in 1973 (Wingert).

Product and Purpose

120. The model produces estimates of operator workload based on the mission time-line. These estimates are intended to validate previous allocation of functions analyses.

When Used

121. The functional interlace model was developed for the analytic evaluation of operator workload as determined by the allocation of functions. It is intended to be used iteratively through concept formulation, evaluation, and selection and design of equipment to satisfy specific operational requirements.

Procedures for Use

122. The user must prepare a detailed mission segment task analysis. An estimate of the performance time of each specific sub-task, or system function, is then assigned to each task, as for TLAP (2.3.2) or WAM (2.3.3). Each task must then be categorised according to its input-output characteristics e.g. 'visual input - motor output', or 'auditory input - no output'. These descriptions permit the user to enter a matrix of interlace coefficients, for each pair of sequential tasks. The times for the two tasks are then sequentially modified by an 'interlace coefficient' which is dependent on the combination of the two operator activities under consideration. Thus:

$$\text{Workload} = (w1 + w2) - I \times w2$$

where w1 and w2 are the task times predicted for functions 1 and 2, and I is the interlace coefficient "typical for the two functions under consideration" (Wingert, 1973). Instead of using the interlace coefficients based on the task input-output characteristics, the user can

generate matrices of interlace coefficients for the different operator functions (i.e. combinations of tasks). These function interlace coefficients can be used at a higher (less detailed) level of system analysis.

Advantages

123. The advantages claimed for the model are that it permits the analysis of continuous, rather than discrete, tasks, and that it produces results which are less conservative than the more typical workload models which treat all sub-tasks as linearly additive.

Limitations

124. As for other time-line analysis models, the functional interlace model requires a very detailed analysis of tasks. Because the interlace estimates are related to input and output processes, the level of detail is fixed, requiring specific operator actions to be identified. If the function interlace approach is to be followed, then each function must have been analysed first at a task or sub-task level. It is not clear from the descriptions available how sequences of multiple task interlaces are handled, i.e. how the interaction of w2 and w3 is dealt with after the interaction of w1 and w2 has been calculated. Pew et al. (1977) suggest that the model appears to require a complete factorial analysis which compares each sub-task with every other sub-task.

Application Examples

125. Wingert (1973) reported comparison of the results with times taken from a simulator study, but stated that the simulation conditions had not been tried out in a full aircraft environment. Pew et al. (1977) reported that the model was still in validation. No other information has been obtained.

Technical Details

126. Wingert (1973) provides a table of interlace coefficients for different input-output conditions. The same reference provides an example of how the interlace coefficients can be combined to describe the extent of interlacing permissible between different operator functions, based on the detailed input-output coefficients. No other technical details are available.

References

Chiles, W.D. (1978). Objective methods. Assessing Pilot Workload. (AGARD-AG-233. Chapter 4). Neuilly-sur-Seine, France. Advisory Group for Aerospace Research and Development.

Jones, A.L., & Wingert, J.W. (1969). On estimating the capability of an avionic man-machine system. International Symposium on Man-Machine Systems. (Vol 2. IEEE Conference Record No 69C58-MMS)

Wingert, J.W., (1973). Function interlace modifications to analytic workload prediction. In K.D. Cross Jr. & J.J. McGrath (Eds.). Crew System Design: An Interagency Conference. Santa Barbara, CA: Anacapa Sciences.

Pew, R.W., Baron, S., Feehrer, C.E., & Miller, D.C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation. (Report No. 3446). Cambridge, MA: Bolt, Beranek, and Newman Inc.

Future Needs

127. Given the potential which Functional Interlacing has to account for time-sharing behaviour, and its close similarity to some concepts of 'attentional demand' (2.3.5) the model appears to warrant further validation and development to incorporate a greater range of input/output processes.

2.3.5 McCracken-Aldrich Model

Summary Description

128. McCracken, Aldrich and their colleagues developed an approach to workload modelling to reflect the mental demands of an operator's tasks. A major goal of the model was to account for concurrent demands on operator attention throughout a system mission. The model is based on a concept of the human operator as an information processor, handling multiple inputs, or 'attentional demands'.

129. Their approach uses a time-line based task analysis in half-second increments, derived from a functional decomposition of system mission segments. Each task is then assigned a rating for operator workload based on the sensory, cognitive, and psychomotor characteristics of the task. The ratings represent difficulty, or effort. The sensory component is further divided into visual, auditory, or kinesthetic inputs. Concurrent ratings are then summed, during each half-second interval of the task time-line.

130. The calculated workload levels are screened using simple rules to identify 'component overload', 'overload condition', 'overload density', and 'system overload'.

History and Source

131. The model originated in the need to conduct workload analyses of different concepts of an advanced, lightweight, multipurpose helicopter, within the space of only six weeks (Aldrich, Craddock, & McCracken, 1984). The original model included seven-point workload scales for visual, cognitive, and psychomotor tasks, and a four-point scale for audio tasks. These workload scales were developed from the ratings of subject matter experts and consist of numerical values and verbal descriptors for each numerical value. Subsequently Szabo and Bierbaum (1986) added a seven-point scale for the kinesthetic and audio sensory components. They also refined the approach to discrete tasks, distinguishing between 'discrete fixed' and 'discrete random' tasks (see Aldrich, Szabo, & Bierbaum, 1989).

132. The model was originally applied through manual analyses of the task data. Subsequently computer programmes were developed to facilitate a more rapid means of analysis, and to permit iterative modifications of the task sequences and attentional demand ratings. In 1988 a model development project was initiated, aimed at validating the attentional demand rating scales, and the overall model, by comparison with expert ratings and the use of a manned simulation facility (Aldrich, Szabo & Bierbaum, 1989).

Product and Purpose

133. The model produces estimates of operator workload in half-second increments of the system missions. Through the application of rules it identifies those combinations of tasks which result in operator 'overload'. This permits identification of the causative factors, for either comparison between different design concepts, or modification and improvement of a specific design.

When Used

134. The model was developed specifically for use in the development and evaluation of system concepts, before specific details of the man-machine interface are available. The emphasis of the model is more on the mission characteristics than the hardware. In its aim, and in subsequent use, the model has proven highly suitable for the cycle of 'function allocation - task analysis - workload analysis' described in the introduction to this chapter. The model has also proven suitable for studying the impact of proposed upgrades to existing aircraft.

Procedures for Use

135. Using the model involves two major phases of effort: the preparation of a mission/task/workload analysis data base, and the compilation and running of the computer-based workload prediction model. The mission analysis is decomposed into mission segments and a list of functions which are performed either sequentially or in parallel. Each function, which is of the typical 'verb - noun' form (e.g. 'control altitude', 'transmit report') is decomposed into the associated aircraft subsystem and operator tasks (e.g. 'detect vertical movement', 'decide if power adjustment needed', 'operate controls'). Each task is assigned a time and a workload rating for the visual, auditory, cognitive, and psychomotor (VACP) components, based on comparison with other tasks, using the scales described above.

136. The computer model uses the task analysis data base, and rules for combining tasks into functions and functions into mission segments, to assemble the overall profile of operator activities. The programme calculates the workload values at one-half second intervals, and prints out when the values change, and when an overload (defined as a total workload rating of 8) occurs.

Advantages

137. The model has the advantage that it can be used early in system development, to explore different system concepts. It includes concurrent tasks, which present difficulties in some other approaches. The main advantage of the model is that it can be used iteratively, either to examine and refine system concepts, or to confirm the estimates of operator workload as the system design becomes more completely detailed.

Limitations

138. The most important limitation of the model is that the assumption of additivity of attentional demands is not proven. This is particularly important in light of the work of Wickens and his colleagues (Derrick & Wickens, 1984) that argues for a partitioning of human encoding (i.e. input) processes. As with others of this class, the model output is heavily dependent on the accuracy of the estimates of operator workload and task times. Other limitations are that it does not treat time stress, and that the task sequence and task times do not vary in response to operator workload. The developers admit that it is labour-intensive to develop initially.

Application Examples

139. The model was used extensively on the U.S. Army LHX light, multipurpose helicopter project. Since development it has also been used to study modifications to the U.S.

Army AH-64A, UH-60A, and CH-47D helicopters (Aldrich, Szabo, & Bierbaum, 1989).

Technical Details

140. Technical details of the computer model are not available at the time of writing.

References

Aldrich, T.B., Craddock, W. & McCracken, J.H. (1984). A computer analysis to predict crew workload during LHX scout-attack missions. (Vol. 1. Technical Report). Fort Rucker, AL: US Army Research Field Unit. MDA903-81-C-0504.

Aldrich, T.B. and Szabo, S.M. (1986). A methodology for predicting crew workload in new weapon systems. Proceedings of Human Factors Society _ 30th Annual Meeting. Santa Monica, CA: Human Factors Society.

Aldrich, T.B., Szabo, S.M., & Bierbaum, C.R. (1989). The development and application of models to predict operator workload during system design. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & E. van Breda (Eds.) Applications of human performance models to system design. New York: Plenum Press.

Szabo, S.M. & Bierbaum, C.R. (1986). A comprehensive task analysis of the AH-64 mission with crew workload estimates and preliminary decision rules for developing an AH-64 workload prediction model. (Draft technical report ASI678-204-86(B), Vols I, II, III, IV). Fort Rucker, AL: Anacapa Sciences Inc.

Derrick, W.L. & Wickens, C.D. (1984). A multiple processing resource explanation of the subjective dimensions of operator workload. (Report EPL-84-2/ONR-84-1). Champaign, IL: Engineering Psychology Research Laboratory, University of Illinois.

Future Needs

141. As indicated in the summary description, a programme of validation and development was proposed in 1988 This included the development of a comprehensive mission/task workload data base, and the capability of varying model parameters, including workload ratings, times, sequences, and extent of automation. The proposals also included the development of an expert system to generate mission functions and segments from the data base.

2.3.6 Attentional Demand Model

Summary Description

142. This model is an extension of the McCracken and Aldrich model (2.3.5), based on the same multiple-resource concept. In lieu of the direct summation of task demands used by McCracken and Aldrich, this model uses 'outcome matrices' to determine the compatibility or interference of multiple attentional demands (visual, audio, cognitive, and psychomotor). The matrices define 'acceptable', 'marginal', and 'unacceptable' paired combinations for each class of attentional demand. A series of decision rules is then applied to determine if segments of operator tasks have acceptable, marginal, or unacceptable workloads.

History and Source

143. Very little has been published on this model. The original work was undertaken at Sikorsky Aircraft Co. by a team which included sub-contractors. The model was validated using man-in-the-loop simulation (R. Peppler, personal communication, September 1986). Subsequently the same sub-contractor used the technique for work with Canadian Marconi Co., who have reported the technique at two international meetings (Makadi 1988). Linton, Plamonden, Dick, Bittner, and Christ (1989) include the model in their state-of-the-art review of operator workload models.

Product and Purpose

144. The model produces plots of 'attentional demand' for visual, audio, cognitive, and psychomotor demands, against the mission time-line. The attentional demands are plotted on a scale of 'acceptable', 'marginal' and 'unacceptable', together with a 'cumulative attentional demand'. Cumulative attentional demand is the sum of the maximum (only) attentional demand ratings across all four demand categories, for all concurrent tasks.

When Used

145. The model was developed for use in concept development, in the cycle of 'function allocation - task analysis - workload analysis' described in the introduction to this chapter. The model can be used once sufficient details of the operator's tasks are available to permit estimates of task times and attentional demands.

Procedures for Use

146. The first step in using the model is to produce a detailed task analysis, to which are added task times and ratings of the visual, auditory, cognitive, and psychomotor demands of each task. The latter are obtained from lists of verbal anchors for each of the attentional categories, using expert judgment.

147. Once the task inventory is complete the outcome matrices are used to screen any simultaneous demands in each attentional category. Six rules are then used to screen the attentional demand profiles, to combine them into an overall rating of 'acceptable', 'marginal', and 'unacceptable'. This includes the application of the 'four second rule', which dictates that

in any condition of simultaneous attentional demand in the cognitive, visual, or psychomotor categories of less than four seconds, the operator will employ time-sharing skills to reduce the workload in the conflicting categories to an acceptable level.

Advantages

148. The advantages of the technique are that it deals with mental workload through the concept of attentional demand, and that it is designed to accommodate time-shared tasks. In addition, the use of the outcome matrices and rules of combination to combine attentional demands avoids the assumption that the ratings are on a ratio scale of measurement. Linton, Plamondon, Dick, Bittner, and Christ (1989) point out that the use of the outcome matrices avoids the interpretive problems which arise, for example, when one class of attentional demand sums to a total of 10 on a scale which ranges from one to seven only. They also note that the model tends to predict slightly higher ratings of workload than those obtained from actual task ratings, which is an advantage for design purposes.

Limitations

149. The main limitation of the technique is that it has not been well documented, and requires additional validation. For example, the use of only the maximum attentional demand ratings from each of the four categories for calculating the cumulative attentional demand can produce intuitively questionable results. This is because the analysis could show that the cumulative attentional demand is acceptable, when the operator is engaged in an undesirably large number of simultaneous tasks. As with other models in this category, it is based on an extensive task inventory, and is therefore dependent on the accuracy of the task analysis and attentional demand ratings (although not overly sensitive to the latter).

Application Examples

150. The technique was developed at Sikorsky Aircraft, and used to study concepts for the LHX light, multipurpose helicopter project. The model was validated using man-in-the-loop simulation. Linton, Plamondon, Dick, Bittner, and Christ (1989) report that the validation showed the model to be sensitive to task differences, and to reflect pilots' subjective workload ratings. The model has also been used to study the workload of the two tactical crew, and of the two pilots, for the proposed Canadian New Shipboard Aircraft (a replacement for the ASW Sea King), for the Canadian light tactical helicopter (CFLH), and for the CF Aurora maritime patrol aircraft update.

Technical Details

151. The model has been set up and run on a VAX 11/750, using the SAINT simulation language as a file system for the task inventory.

References

Hamilton, B.E. & Harper, H.P (1984). Analytic methods for LHX mission and task analysis. Proceedings of Advanced Cockpit Specialist Meeting. Washington, D.C: American Helicopter Society.

Linton, P.M., Flamonden, B.D., Dick, A.O., Bittner, A.C., & Christ, R.E. (1989). Operator workload for military system acquisition. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, K. Sutton, and L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Makadi, J. (1988, May Workload analysis for the Canadian Forces. Poster presentation to NATO workshop on Applications of Human Performance Models to System Design. Ottawa, Canada: Canadian Marconi Co..

Future Needs

152. Because of the size of the task inventory, the model would benefit from the development of an improved data base. A relational data base is being used for the most recent Canadian Forces applications. Current studies are also aimed at comparing the output of the model with subjective workload estimates produced by operators flying actual aircraft missions, or reviewing video-tape records of such missions.

2.3.7 Workload Index (W/INDEX)

Summary Description

153. W/INDEX is a model which is intended to deal with both the temporal and mental demands of a flying task. Through the combination of task time-line analyses and concepts of attentional demand, the model computes the workload demand for each half-second of a mission, by assessing individual task difficulties and the interactions of time-shared attentional demands (North 1986a, 1986b, North & Riley, 1989).

History and Source

154. The model was developed by North at Honeywell Inc. following a series of contracts which required the prediction of operator workload in aircraft crew-stations. The model was originally developed in 1983, and refined in 1984 with the development of the matrix of conflicts between simultaneous attentional demands.

155. Since then additional developments have been undertaken, (North & Riley, 1989), aimed at improving the technique and integrating it with a systematic design process being developed by Honeywell. In 1987/88 the model was being offered as a disc that would run on a IBM PC with 128K memory.

Product and Purpose

156. W/INDEX is intended to predict aircrew workload for a specified mission, identify excessive workload peaks, and identify the factors contributing to the excessive workload. The main aim of the development was to facilitate studies of the impact of automation on aircrew workload, and permit comparison of competing design concepts.

When Used

157. W/INDEX is intended for use early in the development of a crew-station concept, once the sequence of operator tasks has been established. It has been designed to permit use at a gross level of detail, and to permit progressive refinement as the design is defined in detail, and to facilitate iteration of the workload estimates throughout the development cycle.

Procedures for Use

158. The user must generate, or have available, three sets of data:

(1) a task time line (the Operator Activity Timelines, North 1986b);

(2) an interface/activity matrix (generated from the Human Activity Matrix, and the Crewstation Interface Channels list, North 1986b);

(3) an interface/conflict matrix.

159. The task time-line data are developed from a detailed mission analysis. They must include the sequence and start and stop times of each task.

160. The interface/activity data are derived from the task time-lines, and from the human/machine interface channels. The latter are categories of input and output information. Early in concept development, basic channels such as 'visual', 'auditory', 'manual' and 'verbal' can be used. Once the crew-station design has been developed, more specific channels can be used, such as 'windows', 'helmet', 'keyboard'. The interface/activity matrix is completed with the assignment to each cell of a rating of attentional demand for each task/channel combination, on a 1 - 5 scale. These ratings must be developed by the model user, no set of verbal anchors or descriptors is provided.

161. The interface/conflict matrix was developed from studies of human time-sharing behaviour. It consists of values from 0 to 1, assigned on the basis of attentional demand resource conflict, e.g. visual-visual, visual-auditory, visual-manual etc. The values represent conflict between attentional demands: 'high conflict' - 0.7 - 0.9; 'medium conflict' - 0.4 - 0.6; 'low conflict' - 0.2 - 0.4. When applying these values the model user must adjust them, within the given range, for specific factors in the interface such as physical separation or proximity of controls and displays.

162. Once the data are loaded in the model, it calculates operator workload on an instantaneous (half-second) and five-second average basis, where

workload=sum of channel/interface demands + sum of channel/interface conflict penalties

Advantages

163. W/INDEX has the advantage that it can be used early in system development, before the crew-station has been defined and modified, and iterated as the crew-station concept is developed. Its particular advantage, compared with other workload models, is its focus on time-sharing behaviour. The conflict matrices address the costs associated with time-sharing behaviour.

Limitations

164. As with the other models reviewed in this chapter, W/INDEX is heavily dependent on a detailed task analysis, including task times and sequences. The original version suffered from the other common problem of using a fixed sequence of tasks, and using ratings of attentional demand which require additional validation. North (1986a) also notes that the model is insensitive to transient effects such as fatigue or stress. This, again, is common to most of the models reviewed in this chapter.

Application Examples

165. North and Riley (1989) report that the model has been applied to a wide range of systems and problems. It was used to evaluate early versions of the LHX light, multipurpose helicopter, and concepts of one and two-man crews for the Apache helicopter. The model has also been used in studies for the US Advanced Tactical Fighter, and the National Aerospace Plane.

Technical Details

166. North and Riley (1989) report that the model runs in both MS-DOS and VAX-VMS environments. The PC-based version requires 128K of memory.

References

North, R.A. (1986a). A workload index for iterative crewstation evaluation. Proceedings of the Eight Annual Carmel Workshop: Workload and Training: An Examination of their Interactions.

North, R.A. (1986b). WINDEX. H.M. Fiedler (Ed.) Proceedings of the DoD Workshop on Workload Assessment Techniques and Tools. Newport, RI: US Naval Underwater Systems Center.

North, R.A. and Riley, V.A. (1989). W/INDEX: A predictive model of operator workload. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York:Plenum Press.

Future Needs

167. North and Riley (1989) report five developmental efforts to improve W/INDEX. These include: improvements to its predictive accuracy through more detailed consideration of the conflicts that may arise between tasks; improvements to the model's representation of cognitive activities; improvements to the method for producing task/channel attentional demand ratings; development of a capability to modify tasks sequences as a function of the calculated workload level; and integration of the model with a systematic design process.

CHAPTER 3

SINGLE TASK MODELS

3.1 INTRODUCTION

168. Representing human information processing in dynamic situations (Figure 3.1) is the central issue of the Chapters 3 and 4. Wickens (1987) describes the generic information-processing aspects of human operator tasks in this form: "In any task to which the human is assigned, information must be processed. Events and objects in the world must be perceived and interpreted, and then either responded to immediately or stored in memory for later action. Figure 3.1 provides a representation of human information processing that explicitly labels each of these mental activities. Information received through the senses is first perceived. This process of perceptual recognition involves some match between the sensory information and a representation of the recognised object stored in permanent long-term memory. Once a stimulus is identified, a decision must be made as to what action to take. In this case, a response may be selected immediately, or the information may be maintained for some period of time in working memory. If the latter course of action is chosen, the stored information may either be given a more permanent status in long-term memory, forgotten altogether, or used to generate a response. Once a response is selected, it must be executed, normally through a process of coordinated muscular control, operating somewhat independently of the selection that preceded it. Finally, as indicated in the figure, the consequences of a response normally become available again to perception as feedback. This feedback may either be intrinsic - such as the feeling of the fingers, the sound of a key press, or the sound of one's voice; or extrinsic - such as a light that appears on a video display to acknowledge that a command was received. Feedback of both forms is generally helpful to performance, particularly for the novice, and when it is immediate."

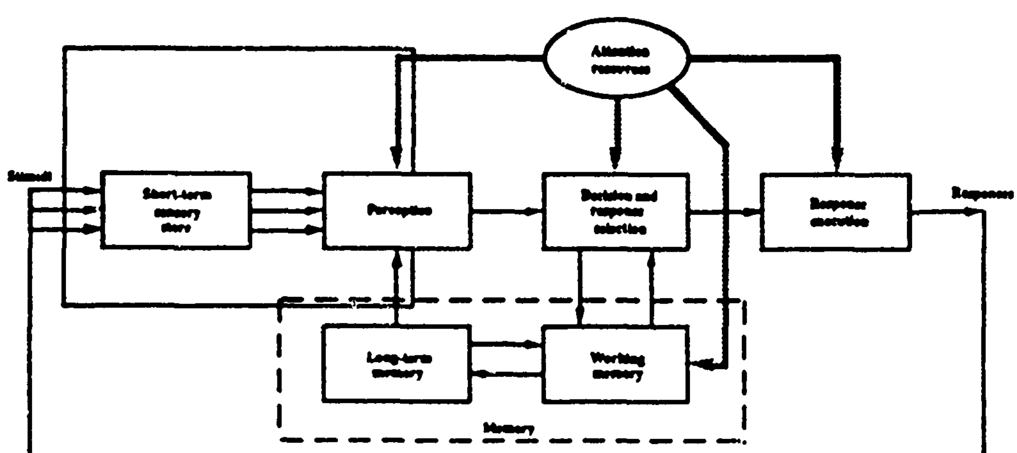


Fig. 3.1 Structure of human information processing (Wickens, 1987).

169. Most of the activities described in Figure 3.1 can take place rapidly, but under constraints reflecting the capacity of the various mental operat-- ɿvolved. These capacities are of two generic forms: (1) Each operation has limits in the ɿ of its functioning and in the amount of information that can be processed in a given un.. ... ɿime. (2) There are limits on the total attention or resources available to the information processing system. These limits are represented by the pool of attentional resources shown at the top of Figure 3.1. This chapter discusses some of the fundamental memory, decisional, response, and attentional limitations of human performance. The approach then focuses on limits of particular mental operations, rather than on characteristics of the entire system. Although it is true that this approach cannot be used exclusively to model performance in more complex settings it is equally true that these more complex modelling efforts must account for the basic limitations and characteristics of the subprocesses. These subprocesses are the concern of the human performance models presented in Chapters 3 and 4 of this report: the relation of displayed and perceived information to its storage in working memory, the limitations of working memory itself, the limitations of cognitive processes related to decision-making and diagnosis, the limitations of response processes as they are manifested in discrete responding tasks, and the limitations and characteristics of attention as they influence the human's ability to carry out two tasks concurrently.

170. In this chapter several single task models (or models of individual tasks) are reviewed. Although many of the models discussed have originated from non-military environments, their contribution and relevance to this report are obvious. Human operator or human performance models are of increasing importance, since an adequate understanding of human performance is essential for the overall success of man-machine systems. One of the main reasons for building models in general and human performance models in particular seems to be that they can serve as an aid to the designer's, scientist's, or user's thinking about the problem being addressed. Human performance modelling arose four decades ago when signal processing and control-theoretic concepts were applied to the task component that is called manual control now. Since then, a continuous development and relevant advances have been achieved in various areas of human performance modelling (Hess, 1987; Kelley, 1968; Knight, 1987; McRuer, Krendel, 1974; Sheridan, Ferrell, 1974; Wickens, 1986).

Definition and Classification of Human Performance Models

171. A model is just a representation of a real system. As such, models can be classified as being either iconic models (e.g., pictures or model cars), analogue models in which a property of the model is substituted for a property of the real system (e.g., slide rules), or mathematical models. The key feature of a mathematical model is the use of symbols, equations, and other mathematical statements to represent reality. Because of the abstract nature of mathematics, mathematical models can be applied to a much greater variety of situations than either iconic or analogue models. This is especially true in decision-making and information-processing situations. Referring to Pew and Baron (1983), a human performance model is a formal, often quantitative, description of the behaviour of one or more people in interaction with equipment. A model of human performance requires first a model or representation of the system and environment with which the people are to function. The ultimate reasons for building models in general, and human performance models in particular, are to provide for:

(1) a systematic framework that reduces the memory load of the investigator, and prompts him not to overlook the important features of the problem;

(2) a basis for extrapolating from the information given to draw new insights and new testable or observable inferences about system or component behaviour;

(3) a system design tool that permits the generation of design solutions directly;

(4) an embodiment of concepts or derived parameters that are useful as measures of performance in the simulated or real environment;

(5) a system component to be used in the operational setting to generate behaviour, for comparison with the actual operator behaviour to anticipate a display of needed data, to introduce alternative strategies or to monitor operator performance;

(6) consideration of otherwise neglected or obscure aspects of the problem.

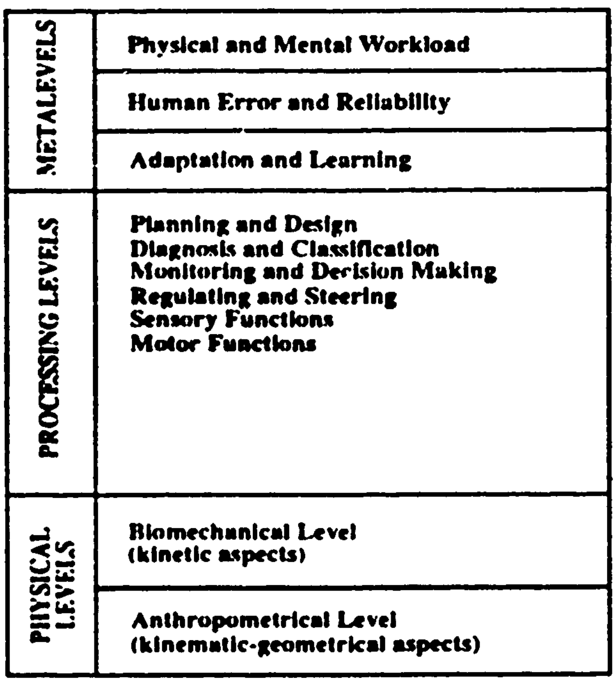


Fig. 3.2 Model-oriented classification of human performance levels.

172. The human factors analyst will experiment with a mathematical model of the system, since it may be impractical to experiment with the real system. This impracticality can result from several factors including:

(1) The real system is only proposed, and therefore, not in existence.

(2) The number and range of feasible values of the design variables are so large that experimenting with the real system would be too costly and too time consuming.

(3) Experimentation with the existing system is unsafe, especially when humans are involved.

173. There exist many different types of mathematical models. These models can usually be classified as being either normative or descriptive, multi-objective or single objective, dynamic or static, and stochastic or deterministic in nature. A normative, or prescriptive, solution is one which optimises the models's objective. One example of such a normative model is the crew model PROCRU in chapter 4. Descriptive models, on the other hand, give the outcome which will occur given particular values for the decision variables. These outcomes can be such quantities as cost, productivity, throughput time, etc. Examples of descriptive models used in human factors research include simulation models and regression equations. Figure 3.2 shows a model-oriented classification of human performance levels.

Levels of Human Performance: The Rasmussen Trichotomy

174. According to Rasmussen (1986), three levels of human performance are identified with very distinct features. The skill-based level represents the highly automated sensori-motor performance which rolls along without much conscious control. The human performs as a multivariable continuous controller for which the functional properties of the system under control are represented as dynamic, spatial patterns. The rule-based performance at the next higher level represents performance based on recognition of situations together with rules

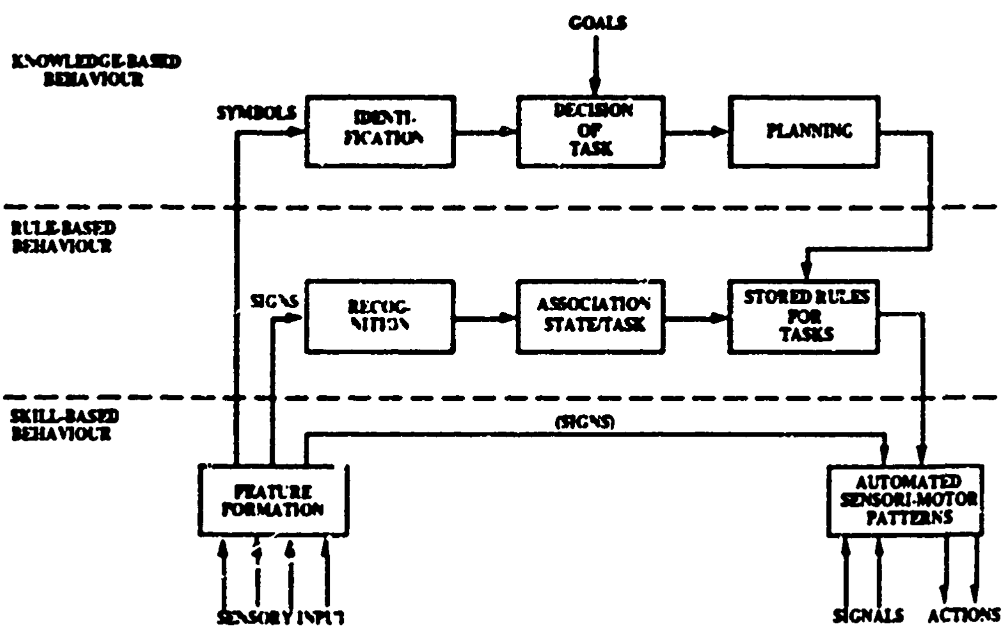


Fig. 3.3 Simplified illustration of three levels of human performance. (Rasmussen, 1986)

for actions from know-how or instructions. The functional properties of the system are at this level implicitly represented by rules relating states and events to actions. The activity at the rule-based level is to coordinate and control a sequence of skilled acts, the size and complexity of which depend on the level of skill in a particular situation.

175. When proper rules and familiar signs are not available for a situation, activity at the next level of knowledge-based performance is necessary to generate a new plan for action ad hoc. The main feature here is that information is perceived as symbols which are used for information processing characterised by an explicit representation of the functional structure of the system to be controlled. The information process used by a person in a specific unfamiliar situation will depend very much on subjective knowledge and preferences and detailed circumstances for the task.

Methodological Position and Validation of Human Performance Models

176. Like simulation and models in science and engineering, human operator models are analogies, which in some way resemble the thing being modelled. A human performance model is not a theory of behaviour. The purpose of a theory is to describe functional relationships, and its value rests on its validity (Chapanis, 1961). Although a model must include functional relationships within its structure, its goal is pragmatic: to predict behaviour and to determine the effects of variables on some system output. Hence, the model is judged on the basis of utility, or the extent to which model outputs assist one to reach a reasonable decision. The extent to which the model represents a nonmodel reality is its validity (Meister, 1985). A conceptual framework for model validity has been developed by Mitchell and Miller (1981).

177. Model validity is one of the most crucial issues in the modelling process and yet, is one of the least understood. According to Meister (1985), the notion of validity is nebulous, rarely defined and poorly understood. One reason may be that model validity is not an independent measure; the validity of a model depends on the purposes of the modelling exercise, the intended uses of the model. Giving a fairly basic definition, model validation is substantiation that a computerised model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model. Even in this case, there are sometimes two views as to what is meant by verification. One view suggests that a model is verified when the model structure and parameters can be adjusted to provide an adequate match to experimental data. The second view holds that a model is validated only when the results of a new experiment are predicted by the model with sufficient accuracy. Many models are useful devices in generating data and exploring bounds on system performance, but have little explanatory power.

178. For either view of model validation, it is necessary to compare experimental results with model predictions and to apply both engineering and formal statistical tests to determine whether or not the model should be considered valid. To accomplish this, one must make judgments concerning (1) the definition of the data, (2) the appropriate figures-of-merit for the engineering and statistical tests, (3) the specific statistical test to use, and (4) the degree of discrepancy between experimental and model results that is considered acceptable. For models of complexity sufficient to represent full-scale man-machine system performance, the problems of validation go well beyond the selection of the proper goodness-of-fit statistics. If

the model is to be useful as a design tool, it must be validated prior to, or at least concurrently with, the development of a full-scale simulation. Due to combinatorial complexity it seems to be nearly impossible to accomplish a full experimental validation of supervisory control and other multi-level models (see Chapter 4).

3.2 OVERVIEW AND RECOMMENDED REFERENCES

179. The models that are available in this category offer the potential user a wide and diversified selection to choose from. Within the limits of the work presented here, it is not possible to completely review such a variety of models, therefore, recourse to the recommended references is suggested for further details if necessary.

180. The model summaries which follow borrow heavily from the original sources given in the lists of references. In many cases direct quotations are not indicated.

3.2.1 References

Ashkenas, I. L. (Ed.). (1988). Advances in flying qualities. AGARD Lecture Series No. 157. Neuilly, France: Advisory Group for Aerospace Research and Development.

Baron, S. & Levison, W. H. (1980). The optimal control model: Status and future issues. Proceedings, IEEE Conf. on Syst., Man, and Cybern. New York: The Institute of Electrical and Electronics Engineers.

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In W. B. Rouse (Ed.), Advances in Man-Machine Systems Research (Vol. 1). Greenwich, CT: JAI Press.

Baron, S. (1987).Application of the optimal control model to assessment of simulator effectiveness. In W. B. Rouse (Ed.). Advances in Man-Machine Systems Research (Vol. 3). Greenwich, CT: JAI Press.

Baron, S. ,Kruser, D. S., & Huey, B. M. (Eds.) (1990). Quantitative modelling of human performance in complex, dynamic systems. Washington, DC: National Academy Press.

Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds.). (1986). Handbook of perception and human performance. New York: J. Wiley

Butler, K., Bennet, J., Polson, P. & Karat, J. (1989). Report on the workshop on analytical models. ACM-SIGCHI Bulletin, 20, (4), 63-79.

Card, S. K., Moran, T. P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum.

Card, S. K., Moran, T. P., & Newell, A. (1986). The model human processor: An engineering model of human performance. In K. R. Boff, L. Kaufmann & J. P. Thomas (Eds.). Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Chapanis, A. (1961). Men, machines, and models. American Psychologist, 16, 113-131.

Chapanis, A. (1967). The relevance of laboratory studies to practical situations. Ergonomics, 10, 557-577.

Chapanis, A. (1972). Human engineering tests and evaluations. In H. P. Van Cott & R. G. Kinkade (Eds.). Human engineering guide to equipment design. Washington, DC: US Government Printing Office.

Chubb, G. P., Laughery, K. R., & Pritsker, A. A. B. (1987). Simulating manned systems. In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Courtois, P.-J. (1985). On time and space decomposition of complex structures. Comm. ACM, 28, 590-603.

Drury, C. G. (1983). Task analysis methods in industry. Applied Ergonomics, 14, 19-28.

Dutton, J. M. & Starbuck, W. H. (1971). Computer simulation of human behavior. New York: J. Wiley.

Dutton, J. M. & Starbuck, W. H. (1971). Computer simulation of human behavior: A history of an intellectual technology. IEEE Trans. on Systems, Man, and Cybernetics, 1, (2), 128 - 171.

Elkind, J. I., Card, S. K., Hochberg, J., & Huey, B. M. (Eds.) (1989). Human performance models for computer-aided engineering. Washington, DC: National Academy Press.

Fleishman, E. A., Quaintance, M. K., & Broedling, L. A. (1984). Taxonomies of human performance. London: Academic Press.

Frost, G. (1972). Man-machine dynamics. In: H. P. Van Cott & R. G. Kinkade (Eds.). Human engineering guide to equipment design. Washington, DC: US Government Printing Office.

Henneman, R. L. (1988): Human problem solving in dynamic environments. In Rouse, W. B. (Ed.). Advances in man-machine systems research. Greenwich, CT: JAI Press.

Hess, R. A. (1987). Feedback control models. In: G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Hollnagel, E., Mancini, G. & Woods, D. D. (Eds.) (1988): Cognitive engineering in complex dynamic worlds. London: Academic Press.

Johannsen, G., Boller, H. E., Donges, E., & Stein, W. (1977). Der Mensch in Regelkreis. Munich: R. Oldenbourg Verlag.

John, B. E. (1988). Contributions to engineering models of human-computer interaction (Report, Dptm. of Psychology. AD-A 195 590). Pittsburgh, PA: Carnegie-Mellon University.

Kelley, C. R. (1968). Manual and automatic control. New York: J. Wiley.

Kieras, D. E. (1988). Towards a practical GOMS model methodology for user interface design. In M. Helander (Ed.). Handbook of human-computer interaction. Amsterdam: North-Holland.

Knight, J. L. (1987). Manual control and tracking. In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Kobsa, A. & Wahlster, W. (Eds.) (1989). User models in dialog systems. Berlin: Springer-Verlag.

McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

McRuer, D. T. & Krendel, E. S. (1974). Mathematical models of human pilot behavior. (AGARD-AG-188). Neuilly, France: AGARD.

McRuer, D. T. (1980). Human dynamics in man-machine systems. Automatica, 16, 237-253.

Meister, D. (1985). Behavioural analysis and measurement methods. New York: J. Wiley.

Mitchell, C. M., & Miller, R. A. (1981). A conceptual framework for model validation. Proceedings, IEEE Conference Cybern. and Society. New York: IEEE.

Moray, N. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufmann & J. P. Thomas (Eds.) Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Patrick. & Duncan, K. D. (Eds.)(1988). Training, human decision making, and control. Amsterdam: North-Holland.

Pew, R. W., Baron, S., Feehrer, C. E. & Miller, D. C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation. (BBN Report No. 3446). Cambridge, MA: BBN Laboratories.

Pew, R. W. & Baron, S (1983) Perspectives on human performance modelling. Automatica, 19, (6), 663-676.

Poulton, E. C. (1974). Tracking skill and manual control. London: Academic Press.

Rasmussen, J. (1986). Information processing and human-machine interaction. An approach to cognitive engineering New York: North-Holland.

Rasmussen, J. (1988): A cognitive engineering approach to the modelling of decision making and its organization in: Process control, emergency management, CAD/CAM, office systems, and library systems. In Rouse, W. B. (Ed.). Advances in man-machine systems research. Greenwich, CT: JAI Press.

Rouse, W. B. (1980). Systems engineering models of human-machine interaction. New York: North Holland.

Rouse, W. B. (1981). Human-computer interaction in the control of dynamic systems. Computing Surveys, 13, (1), 71-99.

Rouse, W. B. & Boff, K. R. (Eds.) (1987). System design. New York: North-Holland.

Salvendy, G. (Ed.) (1987): Handbook of human factors. New York: J. Wiley.

Sheridan, T. B. & Ferrell, W.R. (1974). Man-machine systems: Information, control and decision models of human performance. Cambridge, MA: MIT Press.

Stassen, H. G. (1976). Man as a controller. In: K. F. Kraiss & J. Moraal (Eds.). Introduction to human engineering. Cologne, W. Germany: Verlag TÜV Rheinland.

Stein, W. (1989). Models of human monitoring and decision making in vehicle and process control. McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Wickens, C. D. (1986). The effects of control dynamics on performance. In K. R. Boff, L. Kaufmann & J. P. Thomas (Eds.). Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Wickens, C. D. (1987). Information processing, decision-making, and cognition. In Salvendy, G. (Ed.). Handbook of human factors. New York: J. Wiley.

Young, L. R. (1973). Human control capabilities. In J. F. Parker, V. R. West (Eds.). Bioastronautics data book. Washington, DC: NASA Headquarters.

3.3 MODEL SUMMARIES

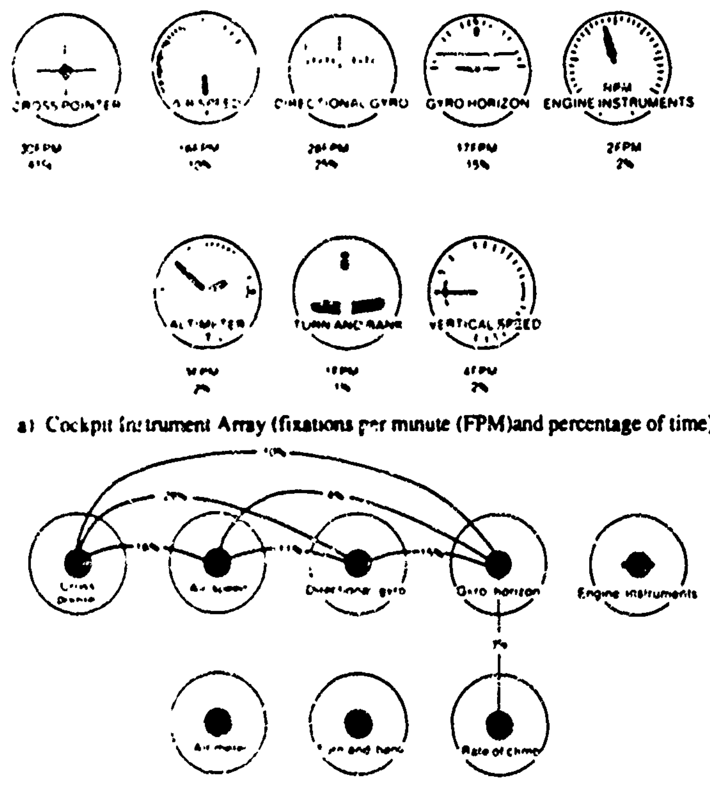
3.3.1 Visual Sampling Models of Senders, Carbonell, and Smallwood

Summary Description

181. The visual sampling model of Senders (1964, 1983) describes the way in which a human operator divides his attention among a number of instruments when he or she scans an array of displayed information. The predictions of the model include, among others, the average duration of a sample, the average percentage of total time that must be devoted to a displayed process, and estimates of the fixation frequencies.

History and Source

182. The first model of human monitoring performance was developed by Senders (1964) in the light of the work of Fitts' group on the eye movements of pilots in cockpits (see Moray, 1986). The model assumes that the spectral characteristics of the displayed signal (e.g., represented in terms of the bandwidths) are the only determinant of the monitor's behaviour. The subsequent models are increasingly complex and take into account more and more factors of real-world tasks such as subjective value and the meaning of the displayed variables.



a) Cockpit Instrument Array (fixations per minute (FPM)and percentage of time)

b) Eye movement Link Values(percentage of transitions,values<2% omitted)

Fig. 3.4 Instrumentation fixations and eye movements of a pilot during a landing approach. (Moray, 1986; after Fitts, 1950).

183. The beginning of investigating human monitoring behaviour (see Figure 3.4) is marked by the work of P. M. Fitts and his group (reviewed by Moray, 1986). These studies included instrument landing approaches of about 40 pilots. The point of fixation was identified by means of camera recordings, but the state of the cockpit instruments was not recorded simultaneously. Figure 3.4 a) shows the positions of the cockpit instruments together with two statistics: FPM is the number of fixations per minute made on that instrument; the second number is the percentage of time during the flight that the pilot spent looking at the instrument. The eye movement link values shown in Figure 3.4 b) are not true transition probabilities. They are the fraction of all eye transitions that went between the instruments indicated. Values less than 2 per cent are omitted. The results of these studies show for example (see Moray, 1986) that the mean fixation duration was 0.6 sec, with a standard deviation of 0.12. The range of fixation durations was 0.4 - 1.4 sec and depended both on the individual pilot and on the instrument that was fixated.

184. According to Moray (1986), a human operator is monitoring a system when he or she scans an array of displayed information without taking any action to change the system state. The purpose of monitoring is to update the operator's knowledge and so to permit appropriate decisions. Monitoring is normally dominated by vision, but auditory signals and communication may be involved, especially when coordinating action. Human decision making as a discrete control activity in the context of man-machine systems may be defined as the process of selecting an appropriate alternative from a set of possible alternatives, based on the perception of actual system states and other sources of information (Sheridan, Ferrell, 1974). Many types of human activity in man-machine systems have decision making as an implicit component, although they would be more commonly considered as sensory-motor, or even cognitive tasks. During phases of pure monitoring or scanning, the nature of actions executed as a result of information obtained by scanning cannot be taken into account. Following these definitions of human monitoring and decision making, a model-oriented classification of related operator tasks can be proposed:

(1) Pure monitoring (i.e., observing without acting).

(2) Independent decisions (binary or multi-valued) based on non-sequential observations.

(3) Independent decisions (binary or multi-valued) based on sequential observations.

(4) Dependent (or dynamic) decisions based on heterogeneous types of observations.

(5) Heterogeneous types of observations and decisions, often embedded in sequences of other task components (e.g., supervisory control situations).

Product and Purpose

185. The model, based upon information theoretic concepts developed by Shannon and others, describes the way in which an operator divides his attention among a number of

instruments while he monitors them. Fundamentally, it assumes that a human operator's fixation frequency for a particular instrument depends upon its information generation rate,

$$\dot{H}_i = W_i \log_2 (A_i^2 / E_i^2) \text{ bits/sec,} \quad (1)$$

where W_i is the bandwidth of the i-th displayed signal, A_i is its amplitude, and E_i is the permissible rms reading error of instrument i. For an observer to reconstruct the signal, the sampling theorem requires that his sampling rate F_i be at least $2 W_i$. If $F_i = 2 W_i$, then the average information to be assimilated by the operator at each fixation is

$$\dot{H}_i/(2W_i) = H_i + \log_2 (A_i/E_i) \text{ bits.} \quad (2)$$

186. For an observer with a fixed channel capacity, who must share his attention among several displays presenting uncorrelated stochastic processes with known information rates, the attentional demand of a particular instrument is calculated to be

$$T_i = 2K W_i \log_2 (A_i/E_i) + 2W_i C \text{ sec/sec,} \quad (3)$$

where T_i is the percentage of total time that must be devoted to displayed process i, K is a constant with dimensions of time per bit, and C (with dimensions of time per fixation) is a constant that accounts for movement time and minimum fixation time. Hence, the duration of a sample is given by

$$D_i = K \log_2 (A_i/E_i) + C \text{ sec.} \quad (4)$$

187. Validation studies, using an array of instruments displaying signals of various amplitudes and bandwidths, showed a good agreement with the model predictions. Values for the constants K and C must be extracted from data collected in a specific context. The model yields estimates of the fixation frequencies and durations for each signal, and for the probabilities of transitions between signals. Limitations of the model became apparent, when an attempt was made to take into account (1) correlations between displayed signals and (2) the interactions between control behaviour and visual sampling.

188. In a recent monograph, Senders (1983) developed his pioneering approach in several ways, provided new and simpler mathematical derivations, and proposed various models depending upon diverse definitions of the goals of the human monitor. The most elementary goal and perhaps the most unreal is that of signal reconstruction. An equally unrealistic goal is that of sampling in the way of pure random choice. Based on these considerations, the following strategies in visual sampling behaviour have been identified that should be regarded as complementary rather than alternatives:

(1) Periodic sampling.

(2) Random constrained sampling.

(3) Conditional sampling:
   (a) Sampling when probability is maximum.
   (b) Sampling when probability exceeds a threshold.
   (c) Variable Nyquist sampling.

(4) Signal reconstruction with imperfect memory.

189. The various models predict different variances for the data. Periodic sampling yields no variability at all, a clearly unrealistic prediction. In the case of random constrained sampling, the variance is a direct consequence of the sampling process. In the case of conditional sampling, the interval is assumed to be a certain mathematical function of the previously observed value.

190. Although each strategy in visual sampling behaviour generates a different mathematical analysis, the assumptions about the signals will be the same:

(1) The signals displayed are random, band-limited time functions with Gaussian amplitude density distributions.

(2) The signals which drive the instruments in an array are assumed to be statistically independent and uncorrelated with one another.

(3) There are always assumed to be three or more instruments in the array.

(4) The different signals displayed do not differ in value.

191. Close to Senders' visual sampling research, Carbonell (1966, 1968) introduced queuing theory concepts to human operator modelling and emphasised the importance of considering the operator's actions. Thus, he moved from the abstraction of Senders' models to more realistic tasks. The human monitor is modelled as a single-channel server that can attend to only one instrument at a time. It is assumed that at each step in his sampling process, the monitor determines for each instrument a subjectively expected cost for not observing it next, and then chooses to observe the instrument with the highest cost of being ignored. An additional assumption is made that the time involved in reading an instrument is constant (approximately 0.33 sec.). According to Pew et al. (1977), the total cost of not looking at any instrument is defined by

$$C(t) = \sum_{i=1}^{M} \frac{c_i P_i(t)}{1 - P_i(t)} \qquad (5)$$

where M is the total number of instruments, c_i is the cost associated with instrument i exceeding its allowed limit, and $P_i(t)$ is the probability that instrument i will exceed its threshold at time t. Thus, the total cost of looking at instrument j at time t is

$$C_j(t) = C(t) - c_j P_j(t). \qquad (6)$$

and the aim of the human monitor will be to choose the instrument j that will minimise $C_j(t)$ at any time.

192. Carbonell's model was compared with Senders' model for eye movement data from realistic landings in an instrument flight simulator using pilots as subjects. It was found to be considerably more accurate than the simpler model but it had to be tuned to each pilot, using his individual estimates of costs, tolerances, and action thresholds. This need to fit the models to individuals emphasises the concept that the operator has his own internal model of the process he is monitoring or controlling. Using Carbonell's model, one must specify the statistical characteristics of each displayed signal, the costs of exceeding given thresholds on each display, and the thresholds below which each instrument reading is ignored. Then the model yields a time sequence of instrument fixations which may be analysed to get visual sampling parameters of interest. A significant feature of Carbonell's model is that the displayed signals are not assumed to be Gaussian with zero mean. Thus, the model represents a significant advance in modelling human sampling behaviour, although it does not attempt to take into consideration cross-coupling among instruments. Referring to Pew et al. (1977), the flexibility and power of this model is obtained at the cost of considerable analytical complexity.

193. Close to Senders' visual sampling research too, Smallwood (1967) developed an instrument monitoring model. His task involved a human operator monitoring the readings of a given number of instruments (e.g., four instruments) which are driven by signals of different amplitudes and bandwidths, and signalling whenever any instrument exceeds a certain threshold. The model assumes that the human operator constructs an internal model of the processes being monitored. The model further assumes (1) that a dead time of about 0.1 second is required to shift attention between two instruments and (2) that the time required to read an instrument inversely related to its distance from the threshold. The predictions of the model include:

(1) Relative fixation frequency for each instrument.

(2) Duration of fixation for each instrument.

(3) Average transition probabilities between the instruments.

194. The concept of the internal model that has been used by Smallwood in an explicit form, plays an important role in human operator behaviour. Indeed human operator models imply an internal model, even when they do not mention it.

When Used

195. The models of Senders, Carbonell, and Smallwood have been developed and validated in various studies. Surveys are given by Moray (1986) and Senders (1983). Principles and components of these models can be embedded into particular approaches and have been successfully applied in many studies. Examples are given by Ellis and Stark (1986), Freund and Sadosky (1967), Kraiss (1981 a, b), Seeberger and Wierwille (1976), and Wierwille (1981).

Procedures for Use

196. To estimate the scanning statistics of the human operator using Sender's model, the bandwidths and amplitudes of the various displayed variables must be specified, along with their acceptable limits. Values for the constants K and C must be deduced from data collected in a specific context. The model yields estimates of the fixation frequencies and durations for each displayed signal, and for the probabilities of transitions between signals.

Advantages

197. The visual sampling research of Senders includes a set of models with increasing complexity for various task situations.

Limitations

198. It is apparent that these models yield good predictions of scanning behaviour under certain circumstances. According to Pew et al. (1977), the failure of these models to account for other results can be attributed to (1) their failure to take into account the redundancy of information obtainable from alternative sets of instruments, which permits controllers to take different scanning strategies under various conditions, and (2) their failure to take into account the interactions between control behaviour and instrument sampling, which permit the controller to estimate changes in the displayed signals. A closely related problem is that particular signals become more or less critical during different manoeuvres, an effect which is largely ignored by these models.

Application Examples

199. A set of experiments was conducted in an aircraft simulator, with pilots flying various types of missions. Eye movement data were recorded along with time histories of the displayed readings on each instrument. Model predictions based upon the amplitude and bandwidth of the various signals proved accurate for some pilots in some phases of flight, but not for all.

Technical Details

200. The visual sampling models of Senders, Carbonell, and Smallwood are well documented, including validation data. The monograph of Senders (1983) gives a comprehensive overview including model refinements and data.

References

Bartlett, M. W. & Smith, L. A. (1973). Design of control and display panels using computer algorithms. Human Factors, 15, (1) 1-7.

Carbonell, J. R. (1966). A queuing model of many-instrument visual sampling. IEEE Trans. Human Factors in Electron., (4), 157-164.

Carbonell, J. R., Ward, J. L. & Senders, J. W. (1968). A queuing model of visual sampling experimental validation. Human Factors in Electron., 9, (3) 82-87.

Ellis, S. R. & Stark, L. (1986). Statistical dependency in visual scanning. Human Factors, 28, (4) 421-438.

Freund, L. E., & Sadosky, T. L. (1967). Linear programming applied to optimization of instrument panel and workplace layout. Human Factors, 9, (4) 295-300.

Gould, J. D. (1968). Visual factors in the design of computer-controlled CRT displays. Human Factors, 10, (4) 359-376.

Kraiss, K.-F. (1981). Analytical evaluation of manned systems with task network models. In J. Moraal & K.-F. Kraiss (Eds.). Manned systems design. New York: Plenum Press.

Kraiss, K.-F. (1981). A display design and evaluation study using task network models. IEEE Trans. Syst., Man, and Cybern., 11, (3) 281-292.

Kvalseth, T. O. (1977). A decision-theoretic model of the sampling behaviour of the human process monitor. IEEE Trans. Syst., Man, and Cybern., 7, (11) 810-813.

Moray, N. (1986). Monitoring behaviour and supervisory control. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.). Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Pew, R. W., Baron, S., Feehrer, C. E., & Miller, D. C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation. (BBN Report 3446). Cambridge, MA: BBN Laboratories.

Seeberger, J. J. & Wierwille, W. W. (1976). Estimating the amount of eye movement data required for panel design and instrument placing. Human Factors, 18, (3) 281-292.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. IEEE Trans. Human Factors in Electron. 5, (1) 2-5.

Senders, J. W. (1983). Visual scanning processes. Tilburg, Netherlands: Tilburg University Press.

Sheridan, T. B. (1970). On how often the supervisor should sample. IEEE Trans. Syst., Sci., and Cybern., 6, 140-145.

Sheridan, T. B. & Ferrell, W. R. (1974). Man-machine systems: Information, control and decision models of human performance. Cambridge, MA: MIT Press.

Smallwood, R. D. (1967). Internal models and the human instrument monitor. IEEE Trans. Human Factors in Electron., 8, (3) 181-187.

Stein, W. (1989). Models of human monitoring and decision making in vehicle and process control. McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Wierwille, W. W. (1981). Statistical techniques for instrument panel arrangement. In J. Moraal & K.-F. Kraiss (Eds.). Manned systems design. New York: Plenum Press.

Future Needs

201. There is interest, and it seems possible, to integrate essential parts of the classical visual sampling approaches (e.g., Senders, Carbonell, and Smallwood) and the control theory approaches to monitoring and decision making (see item 3.3.2 of this chapter).

3.3.2 Control Theory Models of Monitoring and Decision Making

Summary Description

202. Several different monitoring and decision making models have been derived using the information processing structure (i.e., the stages including perceptual limitations, delay, estimator and predictor of Figure 3.6) of the optimal control model (OCM) outlined by Baron (1984), Sheridan and Ferrell (1974), and in other sources. Based on the laboratory paradigm and the related model of Figures 3.5 and 3.6, the following factors of monitoring and decision making have been experimentally studied (including eye-movement recordings) and modelled (Stein, Wewerinke, 1983; Stein, 1989):

(1) Number of displayed processes.

(2) Bandwidths of displayed processes.

(3) Amplitude of processes and probability of displayed events.

(4) Type and intensity of failures embedded in displayed processes (e.g., step, ramp, noise, etc.).

(5) Couplings among unfailed displayed processes.

(6) Couplings among displayed failures.

(7) Size of display array and operator's field of view.

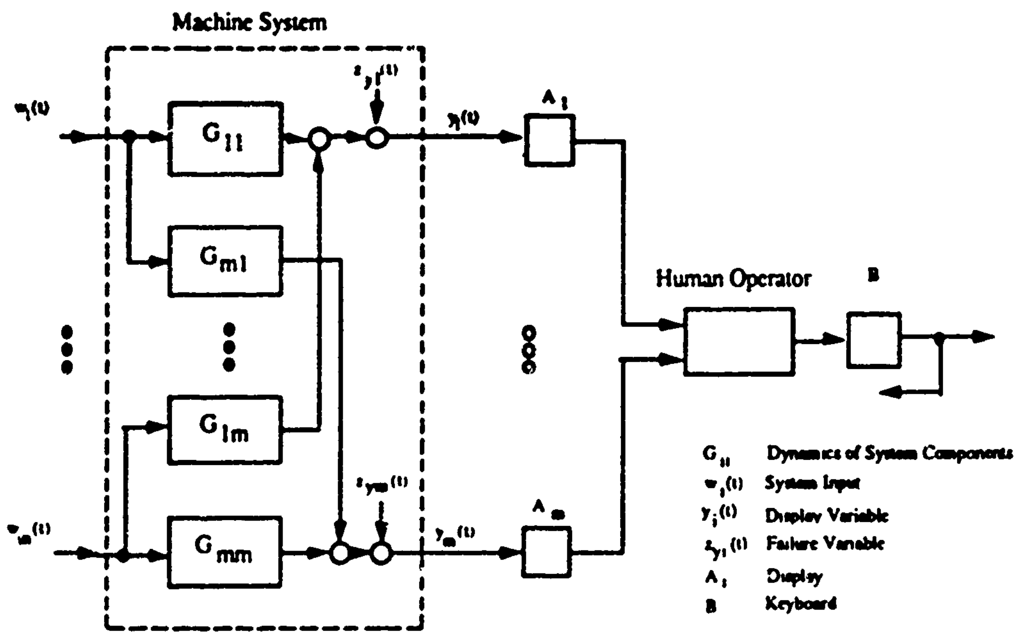


Fig. 3.5 Experimental situation of monitoring and decision making (Stein, Wewerinke, 1983).

History and Source

203. The spectrum of monitoring and decision making tasks and related control-theory models is illustrated in figures 3.5 and 3.6. It is apparent that the estimator-predictor combination of the model produces outputs that can be used for assessing system states and detecting events. Theoretical aspects of these OCM-based monitoring and decision making models are discussed by Phatak and Kleinman (1972) and Kleinman and Curry (1977). Two different types of monitoring and decision making tasks can be found, (1) tolerance-band monitoring and (2) failure detection.

204. A particular type of task including independent decisions and non-sequential observations is tolerance-band monitoring (TBM) which has been extensively studied and modelled by Stein and Wewerinke (1983) and Wewerinke (1976). It is interesting to see that the experimental situations of Senders (1964, 1983), Levison and Tanner (1971), and other researchers (see Moray, 1986) can be characterised as tolerance-band monitoring. TBM tasks involve observing a dynamic process (which can include stochastic and deterministic components) to determine if it is within an explicitly indicated tolerance band. In the binary case, performing a TBM task can be characterised as making independent binary decisions, where each single decision can be based on a single observation testing a pair of hypotheses.

205. A particular type of task including independent decisions and sequential observations is failure detection (FD) which has been extensively studied and modelled by Gai and Curry (1976), Kleinman and Curry (1977), Stein and Wewerinke (1983), and Wewerinke (1983). FD tasks involve observing a dynamic process (which can include stochastic and deterministic components) for the potential occurrence of an abnormal event (e.g., a failure), where an event is defined as a change in the statistics of the displayed process. This change may be constituted for example by changes in mean, standard deviation, dynamic properties, and other characteristics. Performance measures of FD tasks can include speed and accuracy data with related tradeoffs. The detection time denotes the interval between occurrence and detection of a system failure. The accuracy of detecting failures is described by false alarm and miss probabilities. Optimally detecting events or failures with a given accuracy requires sequential observations, i.e., the number of subsequent observations used as input information for making decisions is not fixed, but greater than one. Thus, detecting a failure by observing a displayed dynamic process is a nonstationary binary task that includes testing a pair of hypotheses.

Product and Purpose

206. Based on the information structure of the optimal control model (OCM), three somewhat similar models of monitoring and decision making have been developed:

(1) Monitoring and Decision Making Model of Levison (1971).

(2) Failure Detection Model of Gai and Curry (1976).

(3) Experimental Paradigm and Control Theory Models of Stein and Wewerinke (1983).

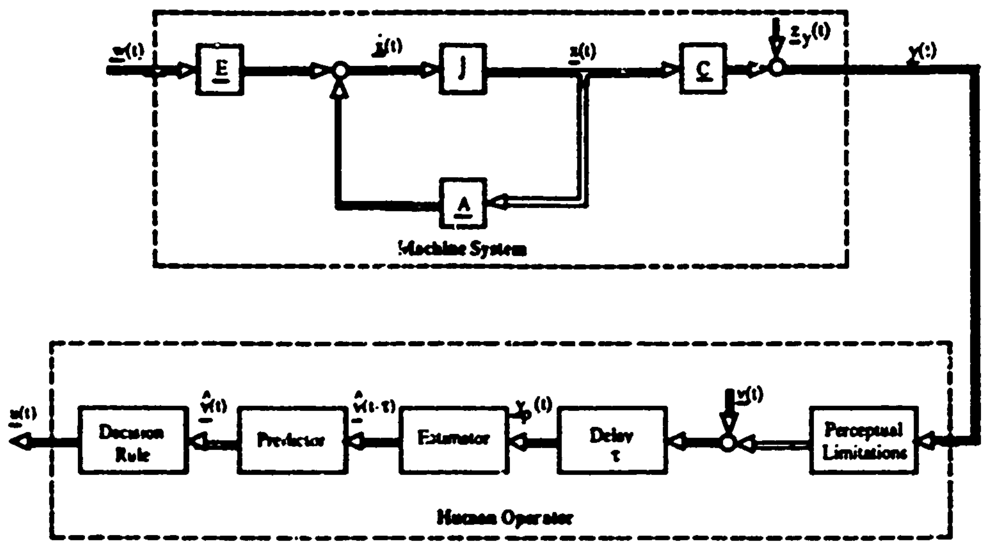


Fig. 3.6 Control theory model of monitoring and decision making (Stein, Wewerinke, 1983).

207. According to Baron (1984), the first use of the OCM information processing structure in modelling monitoring and decision making was by Levison and Tanner (1971). They studied the problem of how well subjects could determine whether a signal embedded in added noise was within specified tolerances. It is assumed that the operator perceives a noisy, delayed version of the displayed processes. The perceived data are then processed, via an optimal estimator-predictor combination, to generate (1) a maximum variance estimate of the system state vector and (2) the covariance of the error in that estimate. This estimator-predictor yield is a sufficient statistic for testing hypotheses about the state of the system. The model assumes that the operator is an optimal decision-maker in the sense of maximising expected utility. This strategy is then applied to the problem of deciding whether or not a signal corrupted by noise, is within certain prescribed tolerances. For equal penalties on missed detections and false alarms, this rule reduces to one of minimising the expected decision error. The resulting decision rule is that of a Bayesian decision maker using a likelihood ratio test. Experimental results have been compared with model predictions for the following task situations: (1) monitoring a single displayed process, (2) monitoring two processes and (3) concurrent manual control and monitoring tasks. Using fixed values for model parameters, model predictions of single-task and two-task decision performance are within an accuracy of 10 per cent.

208. Based on the OCM information processing structure, a failure detection model has been developed by Gai and Curry (1976). They have tested the model in a simple laboratory task and in an experiment simulating pilot monitoring of an automatic approach. In both cases, step or ramp failures were added to an observed signal at a random time to simulate a failure. This produced a non-zero mean value for the signal and for the residual; failure detection consisted of testing an hypothesis concerning the mean of the distribution of the residuals. Sequential analysis was used to perform the hypothesis test. By summing the residuals, a likelihood ratio can be calculated and used to arrive at the decision. Gai and Curry (1976) modified classical sequential analysis to account for the fact that a failure

detection problem is characterised by a transition from one mode of operation to another at a random time, whereas the classical analysis is based on the assumption that the same mode of operation exists during the entire observation interval. They reported good agreement between predicted and observed detection times for both the simple and more realistic situations. In later experiments, the model was used in a multi-instrument monitoring task and accounted for attention sharing and cross-checking of instruments to confirm a failure. A significant result of the experiments was that the property of integration of the residuals appeared to be confirmed for both step and ramp type failures.

209. A laboratory paradigm (Figure 3.5) has been developed by Stein and Wewerinke (1983) as an experimental basis for model-oriented research on various types of monitoring and decision making tasks including eye-movement studies. The corresponding model shown in Figure 3.6, derived from the OCM information processing structure (Baron, 1984), has been developed by Wewerinke (1976, 1983). Thus a highly integrative model of human monitoring and decision making performance is available. By using different decision rules the model can be adapted to different types of tasks:

(1) In the case of independent decisions based on non-sequential observations (e.g., tolerance-band monitoring tasks, TBM), an optimal Bayesian decision rule is involved in the model (Wewerinke, 1976).

(2) In the case of independent decisions based on sequential observations (e.g., failure detection tasks, FD), a sequential decision rule based on a generalised likelihood ratio test is involved in the model (Wewerinke, 1983).

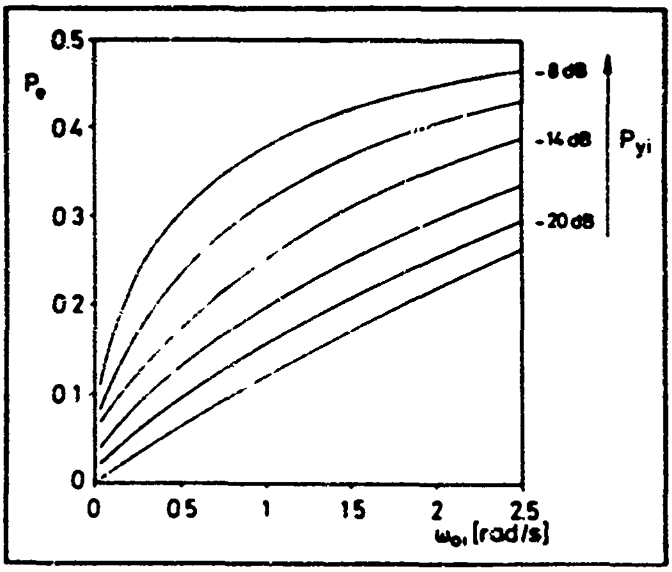


Fig. 3.7 Decision error P_e vs bandwidth ω_{0i} as function of observation noise ratio P_{yi}.

210. An overview of the results of tolerance-band monitoring is given in Figure 3.7 (Stein, Wewerinke, 1983). The correspondence of data and model is high. Human time delay

is assumed to be constant at 0.2 sec. Considering a given task situation with a constant process bandwidth, monitoring performance in terms of the decision error is very sensitive to the observation noise ratio or fraction of human attention devoted to the displayed process; the observation noise ratio increases, when attention is devoted to several processes. The decision error represents the cumulated time fractions of false alarm and missed tolerance-band exceedance. The decision error increases monotonically with bandwidth; the increase begins linearly and becomes progressively nonlinear as a function of both bandwidth and observation noise.

211. An overview of the results of failure de[illegible] is given in Figure 3.8 (Stein, Wewerinke, 1983). These results are restricted to situations with ramp failures. The accuracy of failure detection in terms of false alarm probability is assumed to be constant at a level of 0.05. The detection time increases with observation noise ratio, e.g., when human attention is devoted to several displayed processes. Compared with tolerance-band monitoring, process bandwidth is a factor of minor influence. The predictor portion of the model may be dropped, if human time delay (e.g., 0.2 sec) is small in comparison with detection time.

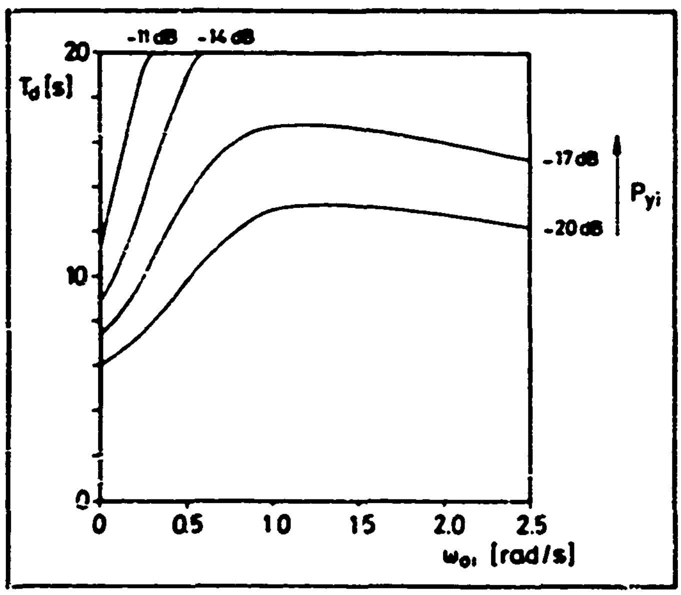


Fig. 3.8 Detection time T_d vs bandwidth ω_{0i} as function of observation noise ratio P_{yi}.

<u>When Used</u>

212. According to Baron and Levison (discussed by Rouse, 1980), the following general display design issues can be addressed using the optimal control model and its derivations:

(1) Is status information acceptable?

(2) Will additional information degrade performance due to interference and/or high workload?

(3) Do the advantages of display integration outweigh the improved scaling possible with separate displays?

(4) Does command information integrate status effectively and, if not, how should it be done?

(5) What performance and workload levels can be achieved with a perfectly integrated and scaled display?

(6) Will quickening, prediction, or preview displays improve performance?

(7) What format should such displays have?

Procedures for Use

213. The application of monitoring and decision making models to the design and evaluation of display arrays and man-machine interfaces can be illustrated in the following way. Given the mathematical equations for the vehicle and process dynamics, the statistical properties of disturbance, and the performance tolerances of each display, the model user is enabled to calculate the fraction of time that the human operator will spend looking at each display, as well as likely transitions among displays. Thus, the model user can determine valuable indices of human operator behaviour. For example, displays that require a relatively large time fraction of looking should be placed near each other, or perhaps be integrated into a single display.

Advantages

214. The advantage of design and evaluation approaches based on the optimal control model and its derivatives stems from (1) the model structure composed of modules for separate human functions (e.g. visual perception, central processing, motor response), (2) the flexible information structure suited for multivariable, multiple process and/or multitask situations (3) the unique performance/workload or performance/attention metric. (4) the comparably high level of model validation, and (5) the underlying normative modelling perspective. The OCM-based approaches are highly developed and seem to be very attractive.

Limitations

215. Model applications are restricted to tasks involving the factors mentioned in the summary description above (e.g., number, amplitudes, and bandwidths of displayed processes, etc.).

Application Examples

216. Design and evaluation studies including the optimal control model and its derivatives have been reported by Baron (1984), Curry, Kleinman, and Hoffman (1977), Moray (1986), and Rouse (1980).

Technical Details

217. Detailed information for understanding and using the monitoring and decision making models is available in the references and the related software documentation.

References

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In W.B. Rouse (Ed.). Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Curry, R.E., Kleinman, D.L., & Hoffman, W.C. (1977). A design procedure for control/display systems. Human Factors, 19, (5) 421-436.

Gai, E.G. & Curry, R.E. (1976). A model of the human observer in failure detection tasks. IEEE Trans. on Syst., Man, and Cybern., 6, (2) 85-94.

Kleinman, D.L. & Curry, R.E. (1977). Some new control theoretic models of human operator display monitoring. IEEE Trans. Syst., Man, and Cybern., 7, (11) 778-784.

Levison, W.H. & Tanner, R. B. (1971). A control-theory model for human decision making. Seventh Annual Conf. Manual Control, (NASA SP-281). Washington, DC: National Aeronautics and Space Administration.

Moray, N. (1986). Monitoring behaviour and supervisory control. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.). Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Pew, R.W., Baron, S., Feehrer, C.E., & Miller, D.C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation. (BBN Report 3446). Cambridge, MA: BBN Laboratories.

Phatak, A.V. & Kleinman, D.L. (1972). Current status of models for the human operator as a controller and decision maker in manned aerospace systems. Proceedings, Automation in Manned Aerospace Systems, (CP-114). Neuilly, France: AGARD.

Rouse, W.B. (1980). Systems engineering models of human-machine interaction. New York: North-Holland.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. IEEE Trans. Human Factors in Electron., 5, (1) 2-5.

Senders, J. W. (1983). Visual scanning processes. Tilburg, Netherlands: Tilburg University Press.

Sheridan, T.B. & Ferrell, W.R. (1974). Man-machine systems: Information, control, and decision models of human performance. Cambridge, MA: MIT Press.

Stein, W. & Wewerinke, P.H. (1983). Human display monitoring and failure detection: control theoretic models and experiments. Automatica, 19, (6) 711-718.

Stein, W. (1989). Models of human monitoring and decision making in vehicle and process control. McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Wewerinke, P.H. (1976). Human control and monitoring - Models and experiments. Proceedings, 12th Ann. Conf. Manual Control. Springfield, VA: National Technical Information Service.

Wewerinke, P.H. (1983). Model of the human observer and decision maker - Theory and validation. Automatica, 19, (6) 693-696.

Future Needs

218. Several needs can be identified, e.g., looking for broader scientific connections between decision making and supervisory control models (see Chapter 4), developing an integration of sampling and decision making approaches, and making model applications more user-friendly.

3.3.3 Auditory Threshold Model

Summary Description

219. This model was developed to predict the hearing threshold in different noise spectra, to permit the design of audio displays (e.g., sonar), audio warning systems and direct voice output (DVO) systems. It is based on work on the human auditory filter, and is a mathematical expression for calculating hearing thresholds across the frequency spectrum of a given noise environment.

History and Source

220. The model is based on a series of experiments conducted to determine the shape of the auditory filter as shown in the references. Patterson argued that, because the rise/fall times of the human auditory systems are short with respect to the duration of speech sounds or signals, and because the relative phase of the spectral components has essentially no effect on masking levels, the auditory threshold in noise can be predicted using a model in which the stimuli are represented by their long-term power spectra and the auditory system is represented by an auditory filter. RAE Farnborough sponsored a joint project with Patterson to develop and evaluate the model (Rood, 1984).

Product and Purpose

221. The model provides estimates of the auditory threshold (dB against frequency), from measurements or predictions of the acoustic noise spectrum at the ear of the listener. The estimates of hearing thresholds can then be used to design auditory displays, by adding 15 dB to the threshold at the signal frequencies.

When Used

222. The model can be used during the development of auditory display systems, hearing protection systems, or the development and evaluation of vehicles and other systems which expose the operators/maintainers to a noise environment which is likely to interfere with communication of one form or another.

Procedures for Use

223. The model is a mathematical formula which has been implemented on a computer at RAE Farnborough. The model requires the characteristics of the noise spectrum as input; other constants have been derived to represent the average listener.

Advantages

224. The model permits the development of auditory displays without the necessity of long and expensive trials to determine hearing thresholds in operational conditions.

Limitations

225. Rood (1984) reports that, in helicopters, at the lowest frequencies, the predicted values of the model are consistently above the measured data. It was concluded that when the dominant component is low frequency rotor noise, the subject listens for the signal in the

throughs between the peak of the masking wave.

226. Patterson has noted that the mathematical expression of the model is applicable to noise levels of up to 95 dB. Above that level the auditory filter broadens, and corrections must be included.

Application Examples

227. Rood (1984) reports an experimental evaluation of the model in which predictions of thresholds based on model parameters derived from the literature were compared with hearing thresholds of ten listeners exposed to simulated noise of Chinook, Sea King and Lynx helicopters. The comparison resulted in a correlation coefficient of 0.990, and a standard error of the estimate of 2.43 dB.

Technical Details

228. The model has been implemented on a Hewlett-Packard Series 300 Microcomputer in BASIC. The frequency characteristics of the noise are determined on a Bruel and Kjaer type 2031/2033 Fast Fourier Transform Analyzer. The data are transferred to the computer via the IEC/IEEE interface by a call routine in the computer programme.

References

Patterson, R.D. (1974). Auditory filter shape. J. Acoust. Soc. Am., 55, 802 - 809.

Patterson, R.D. (1976). Auditory filter shapes derived with noise stimuli. J. Acoust. Soc. Am., 59, 640 - 654.

Patterson, R.D. & Henning, G.B. (1967). Stimulus variability and auditory filter shape. J. Acoust. Soc. Am., 62, 649 - 664

Patterson, R.D. & Nimmo-Smith, I. (1980). Off-frequency listening and audio filter asymmetry. J. Acoust. Soc. Am., 67, 229 - 245.

Rood, G.M. (1984). Predictions of auditory masking in helicopter noise. Paper No. 20, Tenth European Rotorcraft Forum. The Netherlands: The Hague.

Rood, G.M., Patterson, R.D & Lower, M. (1989). An auditory masking model. McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Future Needs

229. Rood (1984) reports that the model will be developed to deal with the detection of signals in low frequency noise, when the listener hears the signal in the wave between noise peaks. No indication is given that the model might be extended to deal with noise levels above 95 dB. Since that level represents the threshold of hazard to hearing for an 8 hour exposure, however, there may be comparatively little use for such a model.

3.3.4 Hick-Hyman Law of Choice Reaction Time

Summary Description

230. The Hick-Hyman law provides a powerful tool for predicting information processing latency of operators confronted with a set of possible events that may occur. The Hick-Hyman law has important implications for system design because the information content of a set of events (warning lights, messages, and so forth) may be defined as important variables that increase the uncertainty of a message set to be responded to. Anything that increases this information content (i.e., increasing the number of possible messages, or varying their relative frequency) can be expected to increase mental workload, increase the chance of errors, and slow the processing time.

History and Source

231. Hick's own experiments used as a display 10 pea-lamps arranged in a somewhat irregular circle. The subject reacted by pressing one of 10 keys on which his fingers rested. The frequencies of the various signals for any given degree of choice were carefully balanced and presented in an irregular order so as to ensure as far as possible that the subject should not be able to predict what signal was coming next.

Product and Purpose

232. This law which relates choice reaction time to the number of choices has been used as the basis for modelling discrete control selection time. In general, reaction time (RT) increases whenever the number of possible stimuli and responses that are appropriate for some situation increases. Simple RT, involving only one stimulus-response pair, is the shortest. In fact, the relation between reaction time and the degree of choice follows a fairly predictable mathematical law known as the Hick-Hyman law (Hick, 1952; Hyman, 1953). The formal expression of the law is

$$RT = a + b \log_2 N \qquad (1)$$

where N is the number of possible equi-likely stimulus-response pairings that could occur in a given context. Because $\log_2 N$ is formally equivalent to the information content of a stimulus in bits, the Hick-Hyman law may be rewritten as

$$RT = a + b\,H_s \qquad (2)$$

where H_s is the information content of the stimulus. In Hick's study the stimuli were 1-10 lights arranged in a near circular display, and the responses were the depression of keys located under the fingers. The value of intercept a was 152 msec and slope b was 111 msec (Keele, 1986). Independently of Hick, Hyman (1953) applied the same formula to data from situations involving one to eight lights assigned verbal responses. Averaged over four subjects, the value of a was 179 msec, and the value of b was 174 msec.

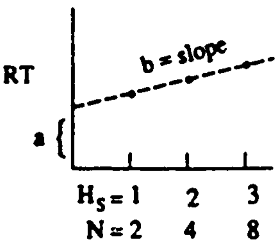


Figure 3.9 Hick-Hyman law: reaction time RT as a function of the number of alternatives N or stimulus information $H_s = \log_2 N$

233. Choice reaction time is affected by the expectancy of the occurring events. A major source of expectancy results from the probability or frequency with which events occur. Probability can also be represented in an informational context because the information conveyed by an event whose probability is p is equal to $\log_2 (1/p)$. The average information conveyed by a series of events with differing probability is simply the weighted average of the individual events' information values. That is,

$$H_{AV} = p_i \log_2 ({}^1/p_i) \tag{3}$$

The Hick-Hyman law has been generalised to estimate mean RT across all events within a set of unequal probabilities. The RT to individual events must be weighted by the probabilities of the events, and the resulting equation is

$$RT = a + b \sum_{i=1}^{N} p_i \log_2 {}^1/p_i \tag{4}$$

where N is the number of possible events and p_i is the probability of an individual event.

234. Choice reaction time is affected by several other factors. The physical relationship or compatibility between a set of stimuli and a set of responses can have a profound influence on the speed of response. Certain compatibility relations are spatially defined. For example, stimuli that are to the right should be responded to with response devices that are also located to the right, and with a rightward movement or clockwise rotation of those devices. Furthermore, physical arrays of stimuli in a certain orientation should preserve the same orientation for their corresponding responses. Where possible, the response made to a stimulus should be physically close to the stimulus itself. Again, large numbers of alternatives increase RT, and stimulus-response (S-R) compatibility decreases RT; the compatibility effect can abolish the effect of number of alternatives. Practice on RT always decreases RT; however, number of alternatives and S-R compatibility interact with practice. That is, RT decreases the most with great incompatibility and a large number of alternatives.

When Used

235. During the design and layout of control/display panels for which minimum operator response time is an important criterion.

Procedures for Use

236. Procedures and data are given in the standard literature (Boff et al., 1986; Boff, Lincoln, 1988).

Advantages

237. The usefulness of information theory in describing human information processing in reaction time is demonstrated because the Hick-Hyman law is found to apply just as well when information is manipulated by probability and context as when it is manipulated by the number of possible stimulus-response pairs (Hyman, 1953; Fitts, Peterson, 1964).

Limitations

238. Several limitations of the Hick-Hyman law have been found (Boff, Lincoln, 1988):

(1) When conditions are not ideal, then RT is more accurately a function of the amount of transmitted information, rather than the number of alternative stimulus-response pairs.

(2) High stimulus-response compatibility can decrease effect of increasing alternatives.

(3) RT results depend on discriminability of alternatives; RT increases as alternatives become more similar.

(4) Mapping multiple stimuli to one response affects RT. For example, RT for four colours (or forms) mapped to two keys (500 msec) is shorter than four colours (or forms) mapped to four keys (572 msec), but is longer than RT for two colours (or forms) mapped to two keys (384 msec).

(5) Choice RT is also affected by stimulus intensity, duration, and probability, as well as by many other factors.

(6) Motor responses to lights occurring with unequal probability yielded shorter RTs overall and less effect of differential probability than for the vocal responses used in this experiment.

Application Examples

239. Applications are surveyed by Boff and Lincoln (1988).

Technical Details

240. See Boff and Lincoln (1988) and the references there.

References

Boff, K. R., Kaufman, L., & Thomas, J. P. (Eds.) (1986). Handbook of perception and human performance (2 Vol.). New York: J. Wiley.

Boff, K. R. & Lincoln, J. E. (1988). Engineering data compendium: human perception and performance (4 Vol.). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Fitts, P. M. & Peterson, J. R. (1964). Information capacity of discrete motor responses. J. Experimental Psychology 67, 103-112.

Hick, W. E. (1952). On the rate of gain of information. Quart. J. Experim. Psychol., 4, 11-26.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. J. Experim. Psychology, 45, 423-432.

Kantowitz, B. H. & Sorkin, R. D. (1983). Human factors: understanding people-system relationships. New York: J. Wiley.

Keele, S. W. (1986). Motor control. In: K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Roe, G. (1982). The heads up, hands back control concept. Proceedings, Conference on Advanced Avionics and the Military Aircraft Man/Machine Interface. (AGARD-CP-329). Neuilly-sur-Seine, France: AGARD.

Salvendy, G. (Ed.) (1987): Handbook of human factors. New York: J. Wiley.

Sheridan, T. B. & Ferrell, W. R. (1974). Man-machine systems: Information, control, and decision models of human performance. Cambridge, MA: MIT Press.

Welford, A.T. (1968). Fundamentals of skill. London: Methuen.

Wickens, C. D. (1987). Information processing, decision making, and cognition. In: G. Salvendy (Ed.), Handbook of human factors. New York: J. Wiley.

Future Needs

241. Studies should be extended considering higher cognitive aspects of human performance.

3.3.5 Fitts' Law: Movement Index of Difficulty

Summary Description

242. The use of Fitts' law for predicting and quantifying operator performance with different control devices has been mentioned by many authors. The law has also been used as a model for predicting reach times to controls in different locations in the work space. The model predicts movement time directly, and deals with accuracy implicitly through the definition of target width W. It has been shown that the basic relationship is valid even when speed of movement is manipulated, as long as effective target width is calculated using the Crossmann correction (Welford, 1968). In most applications, however, the target size is predefined and movement time is the dependent variable.

History and Source

243. The law of movement time has been formulated by Fitts (1954) and seems to describe a basic relationship of human motor processes.

Product and Purpose

244. This empirically derived law predicts the time required to make a speeded, simple positional movement given the distance to be moved and the accuracy constraints to the movement. The following description is mainly based on Boff and Lincoln (1988). For movements in which visual feedback is used (e.g., reaching for an object), movement time MT, which is the interval between initiation of movement and contact with the target (Figure 3.10), is directly related to distance and inversely related to target width (including permissible error tolerance). As Figure 3.11 shows, MT increases with the logarithm of distance (or amplitude) of the movement when target width (accuracy) is fixed, and decreases with the

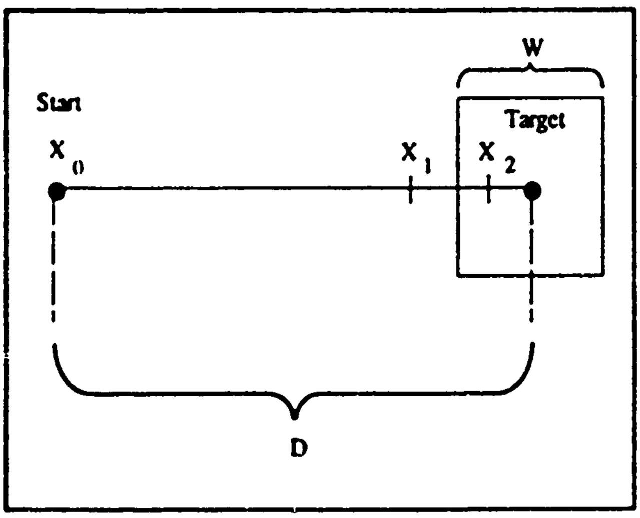


Fig. 3.10 Movement of a user's hand from a starting point over a distance D to a target area width W (Card et al., 1983).

logarithm of target width when distance is fixed. Distance and width are compensatory (i.e., doubling of distance and width produces little change in MT). Fitts' law, which can be used to estimate movement time for movements where accuracy is required, can be stated:

$$MT = a + b \log_2 (2D/W) = a + b\,ID \qquad (1)$$

where a and b are empirically derived constants, D is distance of mov. nt, and W is target width. The constants a and b, representing the intercept and slope of th. ır function, vary over tasks, targets, and subjects. The term $\log_2 (2D/W)$ is sometimes expr..ssed as the index of difficulty ID measured in bits.

245. Fitts' law has to be modified for very short distances. For movement over distances so short that visual feedback cannot be utilised, MT decreases as target size increases, but the function relating MT to log target diameter becomes less linear as movement becomes shorter. This failure of Fitts' law for short movements suggests that short movements are preplanned and that planning for complex movements takes longer. Fitts' law fails for very fast movements that require greater starting and stopping and that do not allow aim correction by visual feedback; this increase in force leads to a decrease in accuracy, which is determined by variability in the preprogrammed muscular impulses. The accuracy of briefer movements is dependent on the speed of movement, (D/MT), and can be described by a variant of Fitts' law, called Schmidt's law:

$$W_e = a + b\,(D/MT) \qquad (2)$$

where a and b are constants, D is movement distance, and MT is movement time. The result of the computation, W_e, is the standard deviation of endpoint dispersion and is known as the effective target width. Schmidt's law provides a good description of accuracy for movements lasting from 140-200 msec over distances of up to 30 cm.

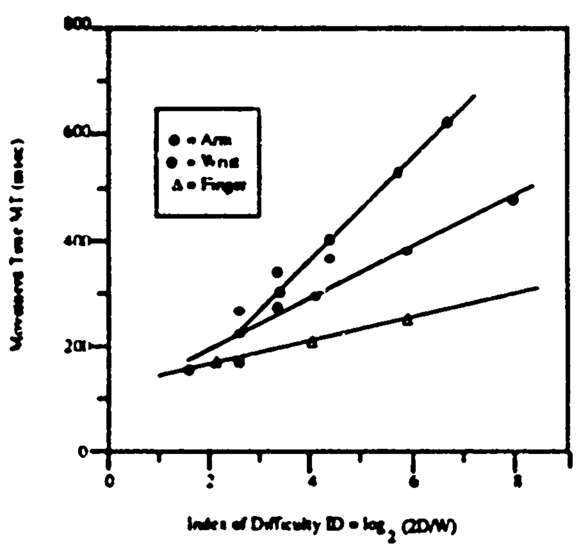


Fig. 3.11 Movement time for finger, wrist, and arm as a function of index of difficulty (Boff et al.,1986).

246. A variant of Fitts' law was developed by Welford (1968),

$$MT = I_M \log_2 (D/W + 0.5), \quad (3)$$

giving a better description of low movement time values than Fitts' basic equation. A value of I_M based on several experiments is set to 100 [50 ~120] msec/bit (Card et al., 1983).

When Used

247. Fitts' law may be used in the design or evaluation of controls and displays for discrete movements such as target acquisition, computer menu selection, etc.

Procedures for Use

248. Applying Fitts' law, movement distance D and target width W are required input parameters. The constants a and b are assumed to be relatively invariant. In each situation the parameters should be checked empirically if the application is to other than simple positional movements of a finger or stylus. In the form shown, the dependent variables is movement time. As discussed above, however, the model could also be used to predict movement accuracy as a function of speed.

Advantages

249. Fitts' law is potentially applicable to any task or subtask in which the layout of positions on a work place is relevant to the design variables under study and where movement time will be a significant proportion of the total time involved in a particular task. The usefulness of the law in the field of human-computer interaction has been demonstrated by Card, Moran, and Newell (1983), since it is the scientific basis for the Keystroke Model and the Model Human Processor.

Limitations

250. Fitts' law is applicable to a wide range of tasks where precise movements are required, but does not apply to movements too brief to permit visual feedback, does not relate MT to reaction time, and offers no description of how visual feedback is used. Further limitations are:

(1) Each application is unique and should be subject to verification, particularly where time-critical performance is required.

(2) When two hands must perform different tasks, MT for the hand performing the easier task (smaller ID) cannot be described by Fitts' law because the harder task (larger ID) determines MT for both hands.

(3) For movements under water, the distance component has a greater effect on MT than does the target width.

Application Examples

251. The Fitts' law methodology has been applied in a large variety of settings, ranging from the control of hand movements under microscopic magnification to positioning of a light spot directed by head movements for a photocell-operated typewriter for paraplegics. There is even some indication that the basic equation is applicable to foot movements, but with different parameter values. It also has been shown to work for positioning pegs into holes and when significant weight is attached to the hand. To a first approximation, it will work for fore-aft movements as well as side-to-side movements, although most available data were obtained for the latter case. It does not predict accurately when the ratio of movement distance to target size is less than about 2 or 1.

Technical Details

252. The formulation of Fitts' law is influenced by information theoretic principles.

References

Boff, K. R. & Lincoln, J. E. (1988). Engineering data compendium: Human perception and performance (4 Vol.). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Beevis, D. (1981). Computer display control devices: A model for predicting performance with different devices (Report IZF 1981-1).Soesterberg, The Netherlands: Institute for Perception.

Card, S. K., Moran, T. P. & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum.

Drury, C. G. (1975). Application of Fitts' law to foot pedal design. Human Factors, 17, 368-373.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. J. Experimental Psychology, 47, 381-391.

Fitts, P. M. & Peterson, J. R. (1964). Information capacity of discrete motor responses. J. Experimental Psychology, 67, 103-112.

Fitts, P. M. & Posner, M. I (1967). Human performance. Belmont, CA: Wadsworth.

Jagacinski, R. J., Repperger, D W., Ward, S. L. & Moran, M. S. (1980). A test of Fitts' law with moving targets. Human factors, 22, 225-233.

Jagacinski, R. J., Repperger, D W., Moran, M. S., Ward, S. L. & Glass, B. (1980). Fitts' law and the microstructure of rapid discrete movements. Journal of Experim. Psychology, 6, 309-320.

Jagacinski, R. J., Plamondon, B D & Miller, R. A. (1987). Describing movement control at two levels of abstraction In Hancock, P. A. (Ed.), Human factors psychology. Amsterdam: North Holland.

Keele, S. W. (1986). Motor control. In: K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Knight, J.L. (1987). Manual control and tracking. In: G. Salvendy (Ed.), Handbook of human factors. New York: J. Wiley.

Langolf, G. D., Chaffin, D. B. & Foulke, J. A. (1976). An investigation of Fitts' law using a wide range of movement amplitudes. J. Motor Behavior, 8, 113-128.

Salvendy, G. (Ed.) (1987): Handbook of human factors. New York: J. Wiley.

Shinar, D. (1986). Human control of robotic movement. Proc., IEEE Conf. on Cybernetics and Society. New York: IEEE.

Welford, A. T. (1968). Fundamentals of skill. London: Methuen

Welford, A. T., Norris, A. H. & Shock, N. W. (1969). Speed and accuracy of movement and their changes with age. In: W. G. Koster (Ed.), Attention and performance II. Amsterdam: North-Holland.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. Acta Psychologica, 41, 67-85.

Future Needs

253. Fitts' law is not restricted to human motor aspects. It has been shown that this speed-accuracy tradeoff can be applied to human control of robot movements (Shinar, 1986) and cognitive aspects of human performance (Wickelgren, 1977).

3.3.6 GOMS Model Methodology for Human-Computer Interaction

Summary Description

254. The GOMS approach (GOMS: goals, operators, methods, and selection rules), developed by Card, Moran, and Newell (1983), is a description, or model, of the knowledge that a user of a computer system must have in order to carry out tasks on a device or system; it is a representation of the "how to do it" knowledge that is required by a system in order to get the intended tasks accomplished. Describing the goals, operators, methods, and selection rules for a set of tasks in a relatively formal way is the goal of doing a GOMS task analysis. The person who is performing such an analysis is referred to as the analyst. Once the GOMS model has been developed, predictions of learning and performance can be obtained as described below. A GOMS description is also a way to characterise a set of design decisions from the point of view of the user, which can make it useful during, as well as after, design. It is also a description of what the user must learn, and so could act as a basis for training and reference documentation. The text of this model summary is based on Kieras (1988).

History and Source

255. The GOMS model notation developed by Card, Moran, and Newell (1983) can be considered as a basis for constructing an explicit model of a computer user's procedural knowledge required by a particular system design. But according to Kieras (1988), there exist several problems in using a cognitive model of a computer user as a design tool. Two critical problems are (1) the difficulty of constructing production rule simulation models; and (2) the difficulty of doing, in a standardised and reliable way, the detailed task analysis required to construct the representation of the procedural knowledge that the user must have in order to operate the system.

256. Primarily for teaching purposes in the field of user interface design and analysis, a guide to GOMS task analysis has been developed by D.E. Kieras (Computer Science Department, University of Michigan, Ann Arbor). The guide defines a language called Natural GOMS Language (NGOMSL) for expressing GOMS models, which has a direct relationship to the underlying production rule models and so supports quantitative predictions and is relatively easy to read and write without knowledge of the production rule models. The guide also includes a procedure for constructing a GOMS model using top-down breadth-first expansion of methods, which seems to be intuitively easier than trying to describe goal structures directly, the approach usually taken in cognitive psychology and by Card, Moran, and Newell (1983).

Product and Purpose

257. NGOMSL is an attempt to define a language that will allow GOMS models to be written down with a high degree of precision, but without the syntactic burden of ordinary formal languages, and that is also easy to read rather than cryptic and abbreviated:

(1) A goal is something that the user tries to accomplish. The analyst attempts to identify and represent the goals that typical users will have. A set of goals usually will have a hierarchical arrangement in which accomplishing a goal may require first accomplishing one or more subgoals. A goal description is an action-object pair in the form: <verb noun>, such as "delete word". The verb can

be complicated if necessary to distinguish between methods. Any parameters or modifiers, such as where a to-be-deleted word is located, are represented in the task description.

(2) Operators are actions that the user executes. There is an important difference between goals and operators. Both take an action-object form, such as the goal of revise-document and the operator of press-key. But in a GOMS model, a goal is something to be accomplished, while an operator is just executed. The observable actions through which the user exchanges information with the system or other objects in the environment are the external operators. These include perceptual operators, which read text from a screen, scan the screen to locate the cursor and so forth, and motor operators, such as pressing a key, or moving a mouse. External operators also include interactions with other objects in the environment, such as turning a page in a marked-up manuscript, or finding the next markup on the manuscript. The internal actions performed by the user are the mental operators; they are non-observed and hypothetical, inferred by the theorist or analyst. In the notation system presented here, some mental operators are built in; these primitive operators correspond to the basic mechanisms of the cognitive processor and are based on production rule models.

(3) A method is a sequence of steps that accomplishes a goal. A step in a method typically consists of an external operator, such a pressing a key, or a set of mental operators involved with setting up and accomplishing a subgoal. Much of the work in analysing a user interface consists of specifying the actual steps that users carry out in order to accomplish goals, so describing the methods is the focus of the task analysis.

(4) The purpose of a selection rule is to route control to the appropriate method to accomplish the goal. The general goal should be decomposed into a set of specific goals, one for each method, and a set of mutually exclusive conditions should be described that specify which method should be used in that context. It is assumed that selection rules come in sets. A set of selection rules is associated with a general goal, and consists of several If-Then rules which choose the specific goal to be accomplished. The relationship with the underlying production rule models is very direct.

258. Based on the GOMS approach, a task description describes a generic task in terms of the goal to be accomplished, the situation information required to specify the goal, and the auxiliary information required to accomplish the goal that might be involved in bypassing descriptions of complex processes. A task instance is a description of a specific task, containing specific values for all of the information in a task description.

When Used

259. In performing a GOMS task analysis, the analyst is repeatedly making decisions about how users view the task in terms of their natural goals and how they decompose the task into subtasks, and what the natural steps are in the user's methods. Once the GOMS model analysis is completed, either for an existing system or one under design, it can be used to estimate the quality of the design. As in any evaluation technique, the measures of design quality are easiest to use if there are at least two systems being compared. Several overall checks can be done that make use of qualitative properties of the GOMS model:

(1) Naturalness of the design - Are the goals and subgoals ones that would make sense to a new user of the system, or will the user have to learn a new way of thinking about the task in order to have the goals make sense?

(2) Completeness of the design - Check that there is a method for each goal and subgoal.

(3) Cleanliness of the design - If there is more than one method for accomplishing a goal, is there a clear and easily stated selection rule for choosing the appropriate method? If not, then some of these methods are probably unnecessary.

(4) Consistency of the design - By consistency is meant method consistency. Check to see that similar goals are accomplished by similar methods.

(5) Efficiency of the design - The most important and frequent goals should be accomplished by relatively short and fast-executing methods.

Procedures for Use

260. Constructing a GOMS model for a system that already exists is the easiest case for the analyst because much of the information needed for the GOMS analysis can be obtained from the system itself, its documentation, its designers, and the present users. The user's goals can be determined by considering the actual and intended use of the system; the methods are determined by what actual steps have to be carried out. The analyst's main problem will be to determine whether what users actually do is what the designers intended them to do, and then go on to decide what the users' actual goals and methods are.

Advantages

261. The GOMS model method description is supposed to be a complete description of the procedural knowledge that the user has to know in order to perform tasks using the system. If the methods have been tested for completeness and accuracy, the procedural documentation can be checked against the methods in the GOMS model.

Limitations

262. Many cognitive processes are too difficult to analyse in a practical context. Examples of such processes are reading, problem-solving, figuring out the best wording for a

sentence, finding a bug in a computer programme, and so forth. Sometimes it is better to bypass the analysis of a complex process by simply representing it with a "dummy" or "placeholder" operator. In this way the analyst does not lose sight of the presence of the process, and can determine many things about what influence it might have on the user's performance with a design.

Application Examples

263. The range of tasks that real systems can be used for is usually very large, and there is practically an infinite number of specific task instances that could be defined. In order to carry out a task analysis, it is usually necessary, for practical reasons, to limit consideration to a subset of the possibilities. This choice of what tasks to analyse has to be made intuitively and informally, but clearly, the tasks should span the major methods and facilities in the system; be high-frequency (often performed) tasks; and be important to be performed quickly and accurately. The essence of the simulation idea is simply to describe a GOMS model for the user's knowledge in a form that is actually executable, so that running the model can verify its completeness and correctness. In the Kieras and Polson work, these models were constructed using the production rule formalism, primarily because this is a standard and current theoretical idea for the representation of procedural knowledge. However, experience in the work suggests that writing production rules is a technically difficult task, analogous in many ways to programming in assembler language.

Technical Details

264. The time to learn a set of methods is basically determined by the length of the methods in terms of the number of NGOMSL statements. There may be little relationship between the number of statements that have to be learned and the number of statements that have to be executed. The situation is exactly analogous to an ordinary computer programme. Execution time may be unrelated to programme length. A GOMS model can predict learning and execution times most plausibly if the operators used in the model are the lowest-level, "standard" primitives, such as pressing a key, which it is reasonable to assume that the learner already knows. Thus, the time to learn a method depends only on how long it takes to learn the content and sequence of steps in the method. In contrast, if the GOMS model is written just at the level of high-level operators that the user has to learn how to perform, then a learning time estimate will have little credibility because the learning times for the operators will be relatively large and unknown. A similar argument applies to execution times. In a method written with high-level operators, the operators will have grossly different, unknown, and relatively long execution times. In contrast, the standard primitive external operators and the relatively simple mental operators such as looking at a manuscript, have relatively small, constant, and known execution times.

References

Card, S.K., Moran, T.P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum.

Kieras, D.E. & Polson, P.G. (1985). An approach to the formal analysis of user complexity. Int. J. Man-Machine Studies, 22, 365-394.

Kieras, D.E. (1988). Towards a practical GOMS model methodology for user interface design. In: M. Helander (Ed.), Handbook of human-computer interaction. Amsterdam: North-Holland.

Kieras, D.E. (1989). The role of cognitive simulation models in the development of advanced training and testing systems. In: N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), Diagnostic monitoring of skill and knowledge acquisition. Hillsdale, NJ: L. Erlbaum.

Polson, P.G. (1987). A quantitative model of human-computer interaction. In: J.M. Carroll (Ed.), Interfacing thought: Cognitive aspects of human-computer interaction. Cambridge, MA: Bradford, MIT Press.

Future Needs

265. What is needed is a higher-level language for describing GOMS models, at the level that designers and specialists in human-computer interaction would normally think about the user's task. Less is known about the relationship between GOMS models and mental workload than for the learning and execution times, so these suggestions are rather speculative. One aspect of mental workload is the user's having to keep track of where he or she is in the mental programme of the method hierarchy. Another aspect of mental workload is working memory load; quantifying this requires making working memory use in the methods explicit.

3.3.7 Keystroke-Level Model of Task Execution Time

Summary Description

266. The Keystroke-Level Model (KLM) describes the time it takes an experienced user to perform a task with a given method on an interactive computer system. The model appears to be simple enough, accurate enough, and flexible enough to be applied to practical design and evaluation situations. This model summary is based on Card et al. (1980) and Boff and Lincoln (1988).

History and Source

267. The Keystroke-Level Model (KLM) has been developed by Card, Moran, and Newell (1980, 1983) at Xerox Palo Alto Research Center. The keystroke model is closely related to the GOMS Model Methodology and the Model Human Processor (MHP), which have been developed by the same authors (summarised in this chapter).

Product and Purpose

268. The prediction problem addressed by the keystroke model is as follows:

Given: A task (possibly involving several subtasks); the command language of a system; the motor skill parameters of the user; the response time parameters of the system; the method used for the task.

Predict: The time an expert user will take to execute the task using the system, providing he uses the method without error.

269. Given a large task, such as editing a large document, a user will break it into a series of small, cognitively manageable, quasi-independent task, which are called unit tasks. The task and the interactive system influence the structure of these unit tasks, but unit tasks appear to owe their existence primarily to the memory limits on human cognition. The importance of unit tasks for analysis is that they permit the time to do a large task to be decomposed into the sum of the times to do its constituent unit tasks. For the purpose here, a unit task has two parts: (1) acquisition of the task and (2) execution of the task acquired. During acquisition the user builds a mental representation of the task, and during execution the user calls on the system facilities to accomplish the task. The total time to do a unit task is the sum of the time for these two parts:

$$T_{task} = T_{acquire} + T_{execute} \quad (1)$$

270. The acquisition time for a unit task depends on the characteristics of the larger task situation in which it occurs. In a manuscript interpretation situation, in which unit tasks are read from a marked-up page or from written instructions, it takes about 2 to 3 seconds to acquire each unit task. In a routine design situation, in which unit tasks are generated in the user's mind, it takes about 5 to 30 seconds to acquire each unit task. In a creative composition situation, it can take even longer. The execution of a unit task involves calling the appropriate system commands. This rarely takes over 20 seconds (assuming the system has a reasonably efficient command syntax). If a task requires a longer execution time, the user will likely break it into smaller unit tasks.

271. A method is a sequence of system commands for executing a unit task that forms a well-integrated segment of a user's behaviour. It is characteristic of an expert user that he encounters and that he can quickly (in about a second) choose the appropriate method in any instance. This is what makes expert user behaviour routine, as opposed to novice user behaviour, which is distinctly nonroutine.

272. Methods can be specified at several levels. A user actually knows a method at all its levels, from a general system-independent functional specification, down through the commands in the language of the computer system, to the keystrokes and device manipulations that actually communicate the method to the system. Models can deal with methods defined at any of these levels. The keystroke model adopts one specific level - the keystroke level - to formalise the notion of a method, leaving all the other levels to be treated informally.

Table 3.1 Description of the Operators in the Keystroke Model

OPERATOR	DESCRIPTION	TIME (SEC)
K	Press key or button (includes shift or control keys). Time varies with skill:	
	Best typist (135 WPM)	0.08
	Average typist (55 wpm)	0.20
	Typing complex codes	0.75
	Worst typist	1.20
P	Point with mouse to target on display (follows Fitt's Law, range 0.8-1.5 sec.)	1.0
H	Home-hands-on keyboard (or other device)	0.40
$D(n_d, l_d)$	Draw n_d straight-line segments of total length $l_d$cm (assumes drawing straight lines with a mouse)	$0.9n_d + 0.6l_d$
M	Mentally prepare	1.35
R(t)	Response by the system (only if it causes the user to wait)	t

273. The keystroke model estimates the time required for an expert user to accomplish a given task using a computer system. Task execution time is described in terms of four physical-motor operators (K, P, H, and D), one mental operator (M), and one system response operator (R), which are described in Table 3.1. An encoding method is given for specifying the series of operators in a task prior to applying the equation:

$$T_{execute} = T_K + T_P + T_H + T_D + T_M + T_R \qquad (2)$$

The keystroke model had been validated against eleven systems.

When Used

274. The keystroke model was evaluated by comparing calculated and observed execution times in ten systems using 14 tasks, 28 operators, 1280 user-system-task interactions. The systems included three text editors, three graphics systems, and four executive subsystems (Card et al., 1980, 1983). Model uses include (1) calculating benchmarks for systems; (2) parametric analysis, where predictions are expressed as functions of task variables; and (3) sensitivity analysis, where changes in the predictions are examined as a function of changes in task or model parameters.

Procedures for Use

275. Given a task (involving a sequence of subtasks, the command language of a system, the motor skill parameters of the user, the response time parameters of the system, and the method used for the task), the keystroke model will predict the time an expert user will take to execute the task. An example application is a text-editing task of replacing a five-letter word with another five-letter word, one line below the previous modification. Using typical keystroke operator times and assuming an average typing speed ($T_K = 0.2$ sec) gives these results:

System A: $T_{execute} = 2T_M + 8T_K + 2T_H + T_P = 6.1$ sec (3)

System B: $T_{execute} = 4T_M + 15T_K = 8.4$ sec (4)

Advantages

276. The keystroke model can be viewed as a system design tool. It is quick and easy to use, is analytical, and does not require specialised psychological knowledge.

Limitations

277. The Keystroke-Level Model (KLM) has several restrictions:

(1) The model applies to the behaviour of experienced users, who have lower variability. No metrics are available for low or moderately experienced operators.

(2) The model assumes error-free performance.

(3) Proper task analysis and encoding are prerequisites.

(4) Tasks that require acquisition time (to perceive, read, or interpret displayed information) are not covered directly by the keystroke model.

(5) With highly repetitive tasks, users reduce their mental time below the model's predictions.

(6) The model does not apply to tasks that emphasise mental operations (e.g., composing text).

278. These restrictions are important and must be carefully considered when using the model. Yet the model seems to represent an appropriate idealisation of this aspect of performance and is a flexible tool allowing the system designer to deal systematically with this aspect of behaviour. The Keystroke-Level Model predicts only one aspect of the total user-computer interaction, namely, the time to perform a task. But there are many other important aspects of performance, there are nonexpert users, and there are nonroutine tasks. All of these must be considered by the system designer. Scientific models do not eliminate the design problem, but only help the designer control the different aspects.

Application Examples

279. Example applications have provided evidence for the keystroke model in a wide range of user-computer interactions. Given the method used, the time required for experts to perform a unit task can be predicted to within about 20 percent by a linear function of a small set of operators. This result is powerful in permitting prediction without having to do any measurements of the actual situation and in expressing the prediction as a simple algebraic expression. The basic application - to predict a time for a specific situation by writing down a method and computing the value - has been sufficiently illustrated in the course of an experiment, where such point predictions were made for 32 different tasks involving 10 highly diverse systems.

Technical Details

280. There exists a very broad documentation of the model and related data (Card et al., 1983) so that an interactive model implementation could be developed.

References

Boff, K.R. & Lincoln, J.E (1988). Engineering data compendium: Human perception and performance (4 Vol.). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Card, S.K., English, W.K. & Burr, B.J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. Ergonomics, 21, 601-613.

Card, S.K., Moran, T.P & Newell, A. (1980). The keystroke level model for user performance time with interactive systems. Comm. ACM, 23, 396-410.

Card, S.K., Moran, T.P. & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum.

Future Needs

281. It is possible to formulate more complicated and refined models than the Keystroke-Model by increasing its accuracy or by relaxing some of its serious restrictions (e.g., models that predict methods or that predict errors). One of the great virtues of the keystroke model is that it puts a lower bound on the effectiveness of new proposals. Any new proposal should do better than the keystroke model to merit serious consideration.

3.3.8 The Model Human Processor

Summary Description

282. The Model Human Processor (MHP) approach of Card, Moran, and Newell (1983) allows the system designer to make general predictions mainly about the time it will take to carry out tasks. The primary means by which it is able to do this is by adopting a strongly parametrised representation. The model human processor can be described by a set of memories and processors together with a set of principles of operation. The model human processor comprises three interacting subsystems: (1) the perceptual system, (2) the motor system, and (3) the cognitive system, each with its own memories and processors. The perceptual system consists of sensors and associated buffer memories, the most important buffer memories being a visual image store and an auditory image store to hold the output of the sensory system while it is being symbolically coded. The cognitive system receives symbolically coded information from the sensory image stores in its working memory and uses previously stored information in long-term memory to make decisions about how to respond. The motor system carries out the response. As an approximation, the information processing of the human is described in the model human processor as if there were a separate processor for each subsystem: a perceptual processor, a cognitive processor, and a motor processor. For some tasks (pressing a key in response to a light) the human must behave as a serial processor. For other tasks (typing, reading, simultaneous translation) integrated, parallel operation of the three subsystems is possible in the manner of three pipelined processors: a typist reads one word with the perceptual processor, passing it on to the cognitive processor, while at the same time typing the previous word with the help of the motor processor.

History and Source

283. The Model Human Processor (MPH) has been developed by S. K. Card, T. P. Moran, and A. Newell at Xerox Palo Alto Research Center (PARC).

Product and Purpose

284. The model human processor is illustrated in Figure 3.12. The working memory consists of activated chunks in the long-term memory. Sensory information flows into the working memory through the perceptual processor. Motor programmes are set in motion through activation of chunks in the working memory. The basic principle of operation of the model human processor is the recognise-act cycle of the cognitive processor: On each cycle, the contents of the working memory activate actions associatively linked to them in the long-term memory, which in turn modify the contents of the working memory. The memories and processors of the model human processor are described by a few parameters. The most important parameters of a memory are

μ, the storage capacity

δ, the decay constant

κ, the main code type.

285. The most important parameter of a processor is τ, the cycle time (the time to process a minimum unit of information). There is no separate parameter for memory access time since it is included in the processor cycle time.

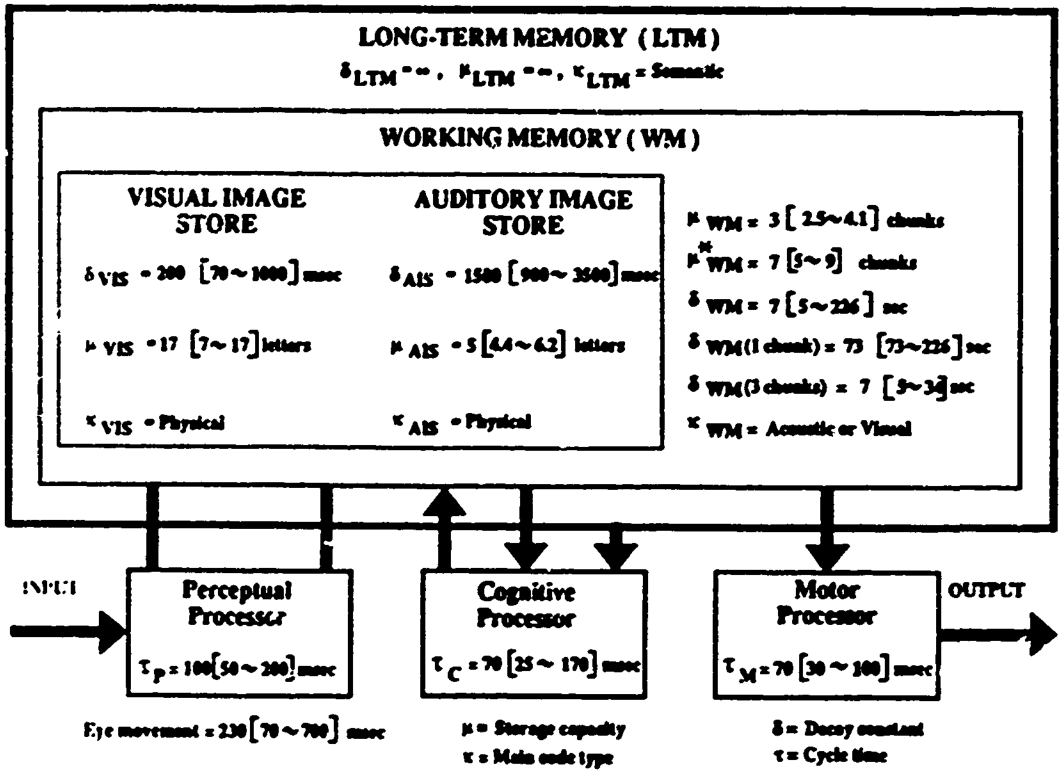


Fig. 3.12 The model human processor (Card et al., 1983).

286. Very shortly after the onset of a visual stimulus, a representation of the stimulus appears in the visual image store of the model human processor. For an auditory stimulus there is a corresponding auditory image store. These sensory memories hold information coded physically, that is, as an unidentified, nonsymbolic analogue to the external stimulus (Figure 3.12):

κ_{VIS} = physical,

κ_{AIS} = physical.

287. The perceptual memories of the model human processor are intimately related to the cognitive working memory as Figure 3.12 depicts schematically. Shortly after a physical representation of a stimulus appears in one of the perceptual memories, a recognised, symbolic, acoustically (or visually) coded representation of at least part of the perceptual memory contents occurs in working memory.

288. As an index of memory decay time, the half-life is used, defined as the time after which the probability of retrieval is less than 50 %. The visual image store has a half-life of about

δ_{VIS} = . (90-1000) msec.

but the auditory image store decays more slowly,

$$\delta_{AIS} = 1500\ (900\text{-}3500)\ \text{msec},$$

consistent with the fact that auditory information must be interpreted over time. The capacity of the visual image store is hard to fix precisely, but for rough working purposes may be taken to be about

$$\mu_{VIS} = 17\ (7\text{-}17)\ \text{letters}.$$

289. The capacity of the auditory image store is even more difficult to fix, but would seem to be around

$$\mu_{AIS} = 5\ (4.4\text{-}6.2)\ \text{letters}.$$

290. The cycle time t_p of the perceptual processor is related to the so-called unit impulse response (the time response of the visual system to a very brief pulse of light), and its duration for purposes of the model human processor is taken to be

$$\tau_p = 100\ (50\text{-}200)\ \text{msec}.$$

291. As modelled by the model human processor, movement is not continuous, but consists of a series of discrete micromovements, each requiring about

$$\tau_M = 70\ (30\text{-}100)\ \text{msec},$$

which is identified as the cycle time of the motor processor. The feedback loop from action to perception is sufficiently long (200-500 msec) that rapid behavioural acts such as typing and speaking must be executed in bursts of preprogrammed motor instructions.

When Used

292. (see Card et al., 1981; 1983; 1986)

Procedures for Use

293. (see Card et al., 1981; 1983; 1986)

Advantages

294. The advantage of the model human processor approach is that it provides a common processing architecture within which a whole class of psychological phenomena can be expressed. One clear advantage of this is that one acquires a common language for characterising a wide range of behavioural data that were previously hard to relate to one another. More importantly, this practice of casting a range of phenomena into a single architecture provides a set of constraints that set theoretically motivated limits on how such behaviours can be described. This stands in contrast to the familiar proliferation of descriptions developed to account for different psychological observations.

Limitations

295. One limitation of the MHP approach is that, in order to achieve its degree of parametrisation, the behaviour that is dealt with must be highly idealised in nature. This is largely due to the fact that since practically no knowledge can be operationally defined at this level of parametrisation there is no capacity to deal with flexibility of behaviour or errors.

Application Examples

296. (see Card et al., 1981; 1983; 1986)

Technical Details

297. (see Card et al., 1981; 1983; 1986)

References

Card, S. K. (1981). The model human processor: A model for making engineering calculations of human performance. Proc. 25th Annual Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society.

Card, S. K., Moran, T. P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum.

Card, S. K., Moran, T. P., & Newell, A. (1986). The model human processor. In: K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of perception and human performance. (Vol. 2). New York: J. Wiley.

Future Needs

298. According to Card et al. (1983), there are at least three areas where the description of the model human processor might be significantly expanded at some cost in simplicity. The first area is the semantic description of long-term memory. As the study of long-term memory proceeded, it became evident to psychologists that, to understand human performance, the semantic organization of long-term memory would have to be taken into account. They have not described semantic memory in any depth here. The second area is the description of the perceptual processor. In the simplified description they have given of perceptual processing, they have skipped over considerable detail that is appropriate at a more refined level of analysis. A description based on Fourier analysis could be used to replace various parts of the model for describing the interactions of visual stimuli with intensity and distance. The third area is the description of the cognitive processor. They have not said much in detail about the control structure of the cognitive processor, but it is necessary to consider the processor's control discipline if interruptability, errors, multiple-tasking, automaticity, and other phenomena are to be thoroughly understood. A more detailed description of the recognise-act cycle, and how the characteristics of simple decisions arise from it, might be given in terms of a set of production rules.

3.3.9 Quasilinear Models and McRuer's Law of Manual Control

Summary Description

299. The quasilinear model approach is used to predict closed-loop system response as a function of the controlled system dynamics and input stimuli. Results are presented in the frequency domain and estimates of phase and gain margins can be deduced. This class of models has been shown to be quite accurate in modelling and predicting performance for compensatory tracking tasks using a single display and controller.

History and Source

300. Although Tustin (1947) was the first to suggest that a servomechanism theory approach could be used to model the human operator in a manual control task, the most significant work undertaken in this area has been by McRuer and coworkers at Systems Technology, Inc., Hawthorne, Calif., USA. This is reflected in the survey publications by McRuer et al. (1967, 1969, 1974) and McRuer (1980) as well as in numerous studies.

Product and Purpose

301. The human controller is complicated to describe quantitatively because of his enormous versatility as an information processing device. Figure 3.13 shows the general pathways required to describe human behaviour in an interactive man-machine system wherein the human operates on visually sensed inputs and communicates with the machine via a manipulative output. This control system block diagram indicates the minimum number of human operator processing stages needed to characterise different behavioural features of the human controller. To describe the components of the figure start at the far right with the controlled element; this is the machine being controlled by the human. To its left is the actual interface between the human and the machine - the neuromuscular actuation system, which is the human's output mechanism. This in itself is a complicated feedback control system capable of operating as an open-loop or combined open-loop/closed-loop system, although that level of complication is not explicit in the simple feedback control system shown here. The neuromuscular system comprises limb, muscle, and manipulator dynamics in the forward loop and muscle spindle and tendon organ ensembles as feedback elements. All these elements operate within the human at the level from the spinal cord to the periphery.

302. There are other sensory sources, such as joint receptors and peripheral vision, which indicate limb output position. These operate through higher centres and are subsumed in the proprioceptive feedback loop incorporating a block at the perceptual level further to the left in the diagram. If motion cues were present, these too could be associated in a proprioceptive-like block. The three other pathways shown at the perceptual level correspond to three different types of control operations on the visually presented system inputs. Depending on which pathway is effectively present, the control structure of the man-machine system can appear to be open-loop, or combination open-loop/closed-loop, or totally closed-loop with respect to visual stimuli. When the compensatory block is appropriate at the perceptual level, the human controller acts in response to errors or controlled element output quantities only. With this pathway operational, continuous closed-loop control is exerted on the machine so as to minimise system errors in the presence of commands and disturbances. Compensatory behaviour will be present when the commands and disturbances are random-appearing and when the only information displayed to the human controller consists of system errors or machine outputs.

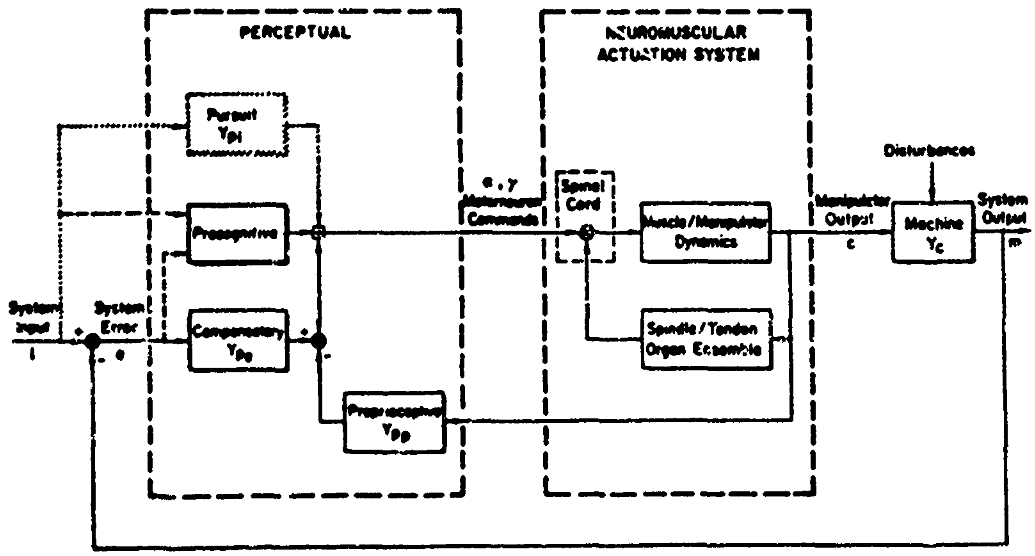


Figure 3.13 Major human operator pathways in a man-machine system (McRuer, 1980).

303. When the command input can be distinguished from the system output by virtue of the display (e.g., signals i and m are shown or detectable as separate entities relative to a reference) or preview (e.g., as in following a curved pathway), the pursuit pathway joins the compensatory. This new pathway provides an open-loop control in conjunction with the compensatory closed-loop error-correcting action. The quality of the overall control can, in principle, be much superior to that where compensatory acts alone. An even higher level of control is possible. When complete familiarity with the controlled element dynamics and the entire perceptual field is achieved, the operator can generate neuromuscular commands which are deft, discrete, properly timed, scaled, and sequenced so as to result in machine outputs which are exactly as desired. These neuromuscular commands are selected from a repertoire of previously learned control movements. They are conditioned responses which may be triggered by the situation and the command and control quantities, but they are not continuously dependent on these quantities. This pure open-loop programmed-control-like behaviour is called precognitive. Like the pursuit pathway, it often appears in company with the compensatory operations as dual-mode control -- a form where the control exerted is initiated and largely accomplished by the precognitive action and then may be completed with compensatory error-reduction operations.

304. A quasilinear or continuous describing function model of the human operator may be defined as the approximate linear model of a non-linear system which minimises the remnant. The remnant is defined as the portion of the human operator's control output power that is not linearly correlated with the system input. Over the years many investigators have

used this basic approach to develop human operator models. However, the two models most widely known and accepted are:

a. The Five Parameter Quasilinear Model:

HUMAN OPERATOR	GAIN	TIME DELAY	SERIES EQUALISATION	NEUROMUSCULAR ACTUATION SYSTEM

$$Y_P(j\omega) = K_P \cdot e^{-j\omega\tau} \cdot \frac{T_L j\omega + 1}{T_I j\omega + 1} \cdot \frac{1}{T_N j\omega + 1} \quad (1)$$

$Y_P(j\omega)$: human operator describing function

T_L/T_I: lead/lag time constants

T_N: neuromuscular time constant

b. The Crossover Model or McRuer's Law:

HUMAN OPERATOR	CONTROLLED ELEMENT

$$Y_P(j\omega) \cdot Y_C(j\omega) = \frac{\omega_c}{j\omega} e^{-j\omega\tau e} \quad (2)$$

$Y_P(j\omega)$: human operator describing function

$Y_C(j\omega)$: controlled element transfer function

ω_c: crossover frequency

τ_e: effective time delay

305. The crossover model or McRuer's law of manual control denotes a relationship in the transfer characteristics of the human operator and the controlled element. The name "crossover" is connected with the model's frequency range of validity. The model assumes that the operator of a dynamic system tries to achieve low error and system stability by behaving in a way that makes the operator and the system together respond as a first-order system in the input bandwidth region. The crossover model describes the product $Y_P \cdot Y_C$ by a first order system with a time delay. The function yields a high gain at low frequencies and a low gain at high frequencies, so the system has low error and is stable. However, there will always be some time delay, τ_e, so there will inevitably be a frequency at which phase lag is greater than 180 deg. According to the crossover model, the operator will maintain a high open-loop gain so that crossover frequency is just below the frequency of 180 deg phase lag and a small phase margin is preserved. The equation is called a describing function rather than a transfer function because the human operator is not truly linear, and the model is thus a quasi-linear model. There is a portion of the human response that is linearly related to system

input and is accounted for by the describing function, but there are also some nonlinear components collectively referred to as remnant. The remnant components contribute a relatively small proportion of the variance in the total response and are added to the output signal. Remnant sources are (1) variations in operator lags not related to input or system dynamics, (2) threshold effects in which small changes are disregarded, (3) intermittency of processing, (4) discrete impulse control, (5) random noise. This is shown in the generalised man-machine system structure of Figure 3.14.

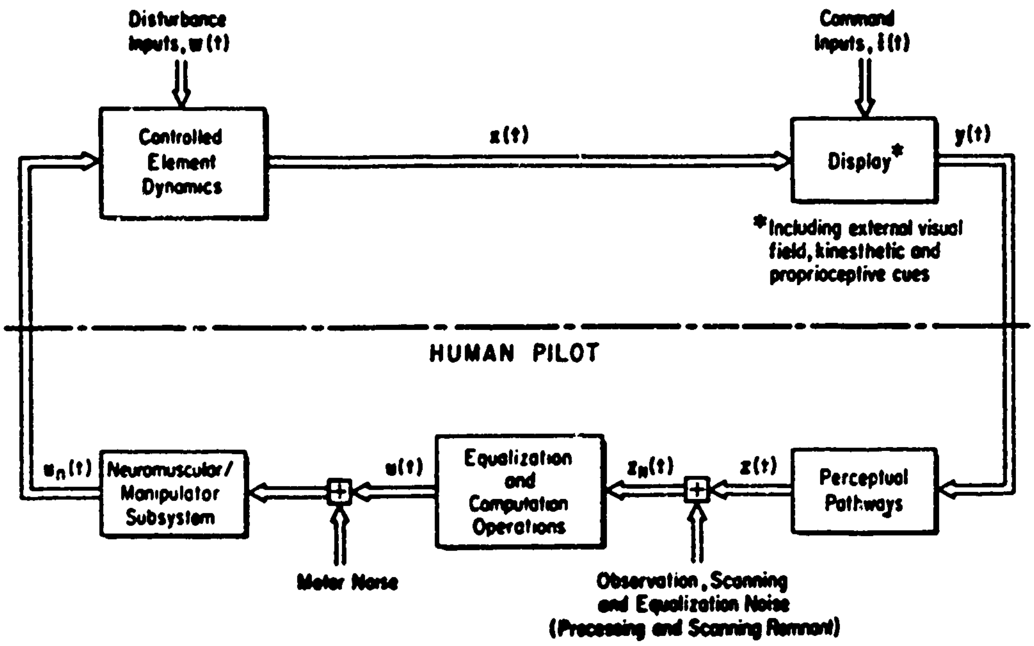


Fig. 3.14 A generalised man-machine system structure (McRuer, 1980).

When Used

306. Stability of aircraft and other control systems can be predicted in the design stage from this model and knowledge of the transfer function of the system. Workload and associated problems such as the effects of fatigue and stress can also be predicted. The models are useful for design of manual control experiments.

Procedures for Use

307. The structure of the models allows a set of free parameters to be adjusted to match the performance for particular task requirements and/or system dynamics. Parameters and their values can be found in the references.

Advantages

308. This approach is a powerful and relatively convenient tool for the analysis of human performance in simple manual control tasks. Validation of the model involves placing human operators in a simulated system and varying parameters such as the transfer function of

the system. Closed-loop stability of aircraft can be predicted from coupling the model with the aircraft transfer function; pilot workload can be predicted from the lead time constant T_L. The model also predicts changes in the Bode plot as a result of operator factors such as stress, fatigue, dual task loading, alcohol, practice, etc. A number of studies seem to bear out the model well.

Limitations

309. The crossover model only applies to a stationary situation, where the task variables are constant and the pilot response characteristics are also stationary and repeatable. The model requires a - 20 dB/decade slope for the combined controlled-element transfer function and operator (describing-function) response. As a quasi-linear model, it exhibits the main features of human operators, but does not model significant nonlinearities, which may be quite large for higher- order systems. As a frequency domain model, it does not easily account for time-domain behaviour such as step responses or transient ramp inputs. As an empirically developed model, it is based on observed human responses rather than on an analysis of the processing mechanisms used by the operator. The effects of mode switching, short-term adaptation, learning of the pilot, or time-varying behaviour in the task variables cannot be treated with this model. Little account is taken of different individual operator styles or strategies. Several limitations are encountered when attempting to model complex cases:

(1) Modelling of multi-input/multi-output tasks can be accomplished, but some expertise is required in specifying loop closures and their specific forms.

(2) The task input signals must be stationary.

(3) Model parameters are likely to vary from task to task thus making this approach less useful for performance prediction.

(4) Operator remnant is generally not dealt with effectively, again limiting accuracy in prediction of man-machine performance.

Application Examples

310. Numerous applications have been cited in the literature, e.g.:

(1) Models for use in car and aircraft control applications (McRuer, Krendel, 1974).

(2) Estimation of vehicle handling qualities (McRuer, Jex, 1967).

(3) Modelling environmental stress effects such as vibration (Jex, Magdaleno, 1978).

(4) Modelling the performance of roll-lateral flight control tasks under visual only or visual-plus-motion situations (Jex et al., 1981).

Technical Details

311. Technical details and information on software (for IBM PCs and compatibles) are given by Allen et al. (1989).

References

Allen, R. W., McRuer, D. T., & Thompson, P. M. (1989). Dynamic systems analysis programmes with classical and optimal control applications of human performance models. McMillan, G. R., Beevis, D., Salas, E., Stuub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Jex, H. R. & Magdaleno, H. R. (1978). Biomechanical models for vibration feed-through to hands and head for a semisupine pilot. Aviation, Space, and Environmental Medicine, 49, 304-317.

Jex, H. R., Magdaleno, H. R., Jewell, W. F., Junker, A. M., & McMillan, G. R. (1981). Effects on target tracking of motion simulator drive-logic filters. (AFAMRL-TR-80-134). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory.

McRuer, D. T., Graham, D., Krendel, E. S., & Reisener, W. (1965). Human pilot dynamics in compensatory systems. (AFFDL-TR-65-15). Wright-Patterson AFB, OH: Flight Dynamics Laboratory.

McRuer, D. T. & Jex, H. R. (1967). A review of quasilinear pilot models. IEEE Trans. Human Factors in Electronics, 8, (3), 231-249.

McRuer, D. T. & Krendel, E. S. (1957). Dynamic response of human operators. (WADC-TR-56-524). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory.

McRuer, D. T & Krendel, E. S. (1974). Mathematical models of human pilot behaviour. (AGARDograph No. 188). Neuilly-sur-Seine: Advisory Group for Aerospace Research and Development.

McRuer, D. T. (1980). Human dynamics in man-machine systems. Automatica, 16, (3), 237-253.

Tustin, A. (1947). The nature of the operator's response in manual control and its implications for controller design. J. Inst. Electr. Engrs., 94, Part II A, 190-202.

McRuer, D. T. & Weir, D. H. (1969). Theory of manual vehicular control. Ergonomics, 12, (4), 599-633.

Future Needs

312. Although developed nearly 30 years ago, the application of the quasilinear models still is in its early stages and many questions remain to be answered. There is need for additional data and appropriate experimental paradigms. Two areas for future research should be mentioned: (1) extension to additional tasks and (2) extensions for skill acquisition. There are many tasks in everyday life that require almost continuous human control for their

successful and safe completion. Driving an automobile, riding a bicycle, and flying an aircraft are three examples among many. Each of these tasks involve the human being acting as a feedback element in a closed-loop control system. Indeed, the control theory paradigm which has evolved in the intervening years has been so useful in quantifying control-related human behaviour that it has become a fundamental mode of thinking on the part of most manual control practitioners. Development of models for learning behaviour (i. e., control-strategy development) is one of the important remaining theoretical frontiers in manual control. Models of this sort have ready application to the design of training simulators.

3.3.10 Optimal Control Model (OCM) of Man-Machine Systems

Summary Description

313. The optimal control model (OCM) represents human performance in manually controlled systems. The model is able to predict task performance (e.g., rms control activity and error) and human control characteristics (e.g., Bode plots of human operator describing functions). What differentiates the optimal control model from other models of the human operator are the methods used to represent human limitations, the inclusion in the model of elements that compensate optimally for these limitations, explicit representation of the human's internalisation ("internal model") of the control task, and the extensive use of state-space concepts and techniques of modern control theory.

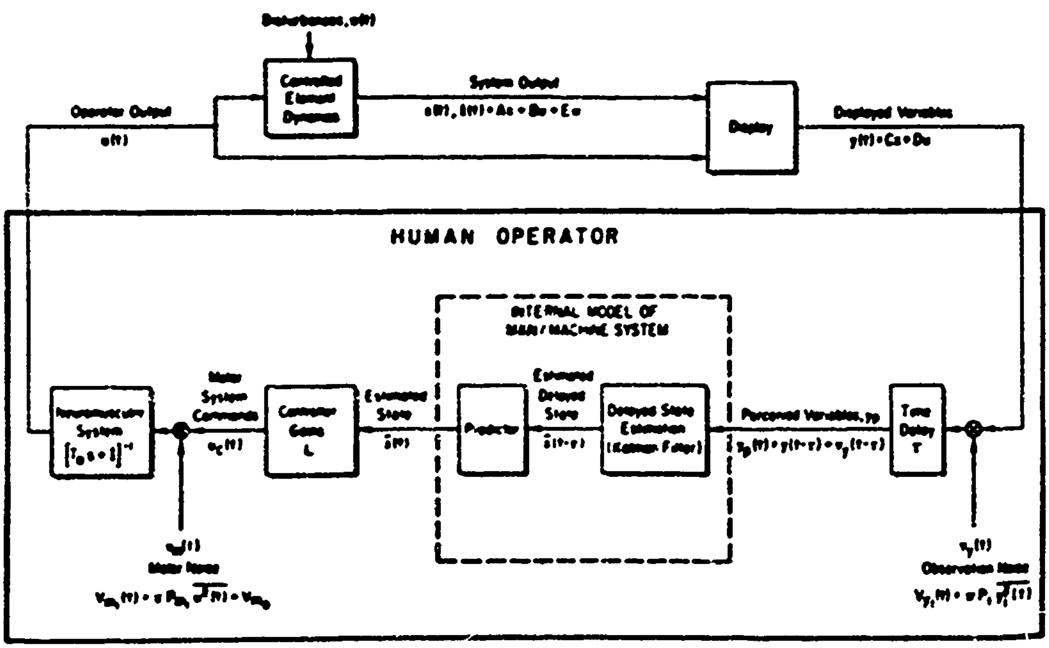


Fig. 3.15 Optimal control model (OCM) of man-machine systems

History and Source

314. The model was developed by Kleinman, Baron, and Levison (1971) at Bolt, Beranek and Newman, Inc., Cambridge, Mass. Surveys are given by Baron and Levison (1980), Baron (1984), and Levison (1989).

Product and Purpose

315. The optimal control model (OCM) has been developed using modern control and estimation theory. It assumes that a well-trained and well-motivated human controller behaves optimally to achieve a specified performance criterion J, subject to certain internal constraints on the human's information-processing capabilities and subject to the operator's understanding of task objectives. This assumption is consistent with notions of human response behaviour discussed in the psychological literature. The model is a stochastic time

domain model of the human. It explicitly models human remnant characteristics and is not limited to stationary inputs. The model further assumes that the human controller has an internal model of both the system and the forcing function which drives it. This model enables an estimation of system state which is limited by available perceptual data that are noisy and delayed. Further limitations include additional motor noise and a neuromuscular lag resulting in a maximum bandwidth. The optimal control model (Figure 3.15) is composed of several submodels:

(1) A perceptual model that translates displayed variables into noisy, delayed, perceived variables.

(2) An information processor consisting of an optimal estimator and a predictor that generates estimates of the system states based on the perceived variables.

(3) A set of optimal gains selected to minimise a chosen quadratic cost function.

(4) An output model that accounts for human motor-response limitations including bandwidth and noise.

316. In the optimal control model (Fig. 3.15), the information from the display is corrupted by observation noise introduced by the human operator. This noisy representation of the display information is then delayed by an amount, τ, representing the internal human processing delay. The model then uses a Kalman filter and predictor to estimate system state. The control motion is then generated with the optimal gain matrix operating on the best estimate, $\hat{x}(t)$, of the system state, $x(t)$. The optimal gain matrix is determined by minimising a quadratic cost functional,

$$J = E\left\{ \lim_{T \to \infty} \frac{1}{T} \int_0^T \left[\sum_i q_i x_i^2 + r_i u_i^2 + g_i \dot{u}_i^2 \right] dt \right\} \tag{1}$$

where x is the system state variable, u and $\dot{u}$ are control variables, and q,r, and g are weightings for the different variables. This reduces to

$$J = E\left\{ \sum_i q_i x_i^2 + r_i u_i^2 + g_i \dot{u}_i^2 \right\} dt \tag{2}$$

in the steady state.

317. Just as an observation noise is postulated to account for perceptual and central processing inadequacies, a motor noise is introduced to account for the human operator's inability to generate noise-free control actions. Finally, the noisy control response is smoothed

by a filter that accounts for an operator bandwidth constraint. In this manner, noise is treated as an integral part of the model, rather than as an external remnant, which is the approach used by the crossover model.

When Used

318. It can be used especially at the design stage in the development of a prototype vehicle. It can be used in the design of vehicle or process control systems, especially those which involve transient (step or ramp) inputs; modelling human decision making, reliability assessment, and workload; design of slowly responding and monitoring systems; design of simulations and experiments. The results can be either in graphical or statistical form.

Procedures for Use

319. Two basic categories of parameters must be adjusted in order to predict performance: (a) parameters which represent inherent human perceptual-motor limitations, and (b) parameters which represent information sources available to the human in a given task. Parameters and values are included in the references. To apply the optimal control model, the following features of the machine system/environment must be specified:

(1) A linearised state variable representation of the system being controlled.

(2) A stochastic or deterministic representation of the driving function of environmental disturbances to the system.

(3) A linearised display vector summarising the sensory information used by the operator.

(4) A quantitative statement of the performance criterion or cost functional for assessing operator/machine performance.

Advantages

320. The state space structure of the model makes it very flexible for use in a variety of control tasks. Transitions from time domain to frequency domain are possible. The model was validated in a set of manual control experiments in which experienced operators tracked systems with K, K/s, and K/s^2 dynamics in a compensatory task (Kleinman, Baron, Levison, 1971). In addition, the OCM was validated against simulated results relating to the hover control of a VTOL aircraft.

Limitations

321. As with all optimal control solutions the OCM is only optimal in the sense of the chosen quadratic cost function and may well be sub-optimal in others. The model describes the behaviour of skilled operators of dynamic systems, but it does not answer the question of how human operators come to behave in an optimal manner. Further restrictions are:

(1) The model is validated for use in continuous systems.

(2) The model is limited to systems which can be linearised.

(3) The application of the model requires considerable experience with its use.

(4) The model is complex, and its use in many applications requires a large amount of computation.

Application Examples

322. The OCM has been applied (mostly with regard to aircraft flight) as a predictive and as a diagnostic tool. Areas of application include display design and evaluation, control design and evaluation, prediction of aircraft handling qualities, simulator design and evaluation, analysis of tasks involving transient manoeuvres, effects of environmental stress, and supervisory control. Examples in each area are summarised in this section.

323. The OCM has been used in numerous studies to evaluate the effectiveness of aircraft flight displays (Baron, Levison, 1975, 1977). In some cases, model predictions have been compared with data; in others, the OCM has been employed as a purely predictive tool. The effects of different display formats and display symbology were predicted in cases where the aircraft was subjected to turbulence and/or constant updrafts. The ability of the pilot to estimate these external disturbances, and take the appropriate corrective action to minimise glide path errors was analysed. Predictions of system performance were compared with data obtained in independent experimental investigations. The model-data agreements were excellent and demonstrated the model's ability to predict the time-varying adaptability of a pilot to updraft disturbances.

324. Although display problems have received the most attention, other aspects of the system design problem have not been neglected completely. Levison and Houck (1975) have explored the use of the model in analysing control stick design problems in a vibration environment. Stengel and Broussard (1978) have used the basic structure of the OCM along with some assumptions concerning sub-optimal adaptation to determine stability boundaries in high-g manoeuvring flight. Schmidt (1979) has proposed a design procedure for stability augmentation systems based on closed-loop analysis with the OCM.

325. Display and control design are two of the factors that influence a vehicle's handling qualities as reflected in the pilot's ability to achieve acceptable system performance at reasonable levels of mental workload. Although aircraft handling qualities are specified, for the most part, in terms of vehicle response characteristics alone, the formal acquisition of subjective pilot opinion is an important aspect of aircraft evaluation and acceptance. Thus, a need exists for a reliable analytic tool for predicting pilot opinion ratings, especially for new aircraft configurations and task environments. Hess (1977) noticed that, for a variety of experimental results that he matched with the OCM, the objective performance index varied monotonically with subjective pilot opinion ratings, and he suggested use of the OCM as a predictor for pilot ratings. Because of the rich set of performance metrics that can be derived from steady-state tracking data, most OCM validation studies (and most applications) have involved consideration of steady-state control tasks. Nevertheless, the OCM is theoretically capable of treating non-steady-state control tasks.

326. Military operational environments may subject the human operator to substantial physical stress. In some cases, the stress is a direct consequence of the flight task (e. g., vibration, sustained high acceleration); in other cases, stress may be induced by an opponent as a defensive measure (e. g., optical countermeasure). Such considerations have motivated application of the OCM to tasks involving actual and simulated environmental stress. A series

of studies was conducted to develop a methodology for modelling the effects of high-frequency vibration on pilot response behaviour and total system performance (Levison, 1978). This effort led to a model structure which combines the OCM with a biodynamic model of the human operator. As part of this structure, a set of rules were developed for relating certain OCM parameters (specifically, time delay and motor noise covariance) to biodynamic response parameters.

327. The above studies all focused on the operator in continuous control tasks. But the structure of the OCM, particularly the information processing sub-model, also lends itself to modelling tasks in which monitoring and decision-making are the major concerns of the operator. The first attempt to exploit this aspect of the OCM was by Levison and Tanner (1971) who studied the problem of how well subjects could determine whether a signal, embedded in added noise, was within specified tolerances. Their experiments were a visual analogue of classical signal detection experiments except that "signal-present" corresponded to the situation of the signal being within tolerance. They retained the estimator/predictor and the equivalent perceptual models of the OCM and replaced the control law with an optimal (Bayesian) decision rule just as has been used in some other behavioural decision-theory models. Phatak and Kleinman (1972) examined the application of the OCM information processing structure to failure detection and suggested several possible theoretical approaches to the problem. Gai and Curry (1976) used the OCM information processing structure to analyse failure detection in a simple laboratory task and in an experiment simulating pilot monitoring of an automatic approach. They reported good agreement between predicted and observed detection times for both the simple and more realistic situations. In the latter case, the model was used in a multi-instrument monitoring task and accounted for attention sharing in the usual OCM fashion.

328. It has also been extensively applied by the US Air Force Aerospace Medical Research Laboratory to a number of manual control problems, e.g.,

(1) Modelling single axis flight control performance under visual only or visual-plus-motion conditions (Levison, Junker, 1977).

(2) Modelling the effects of acceleration stress on human performance during air-to-air tracking tasks (Korn, Kleinman, 1978).

(3) Studying the effects of high frequency vibration (Levison, 1978).

(4) Prediction of altitude holding performance in the presence of linear perspective visual cues (Levison, Warren, 1984).

Technical Details

329. Several implementations of the optimal control model are available. The MANMOD computer programme (Baron, Berliner, 1974) has been developed to implement a flexible and efficient version of the model that can be used to study time-dependent effects, such as disturbance variations and instrument switchovers. In addition, the incorporation of display-related thresholds and resolution limitations allows one to study the effects of instrument modification. MANMOD (written in FORTRAN) has been designed for interactive operation and can be extended to a rather general programme for a wide field of applications, including display monitoring and decision making, in stationary and unstationary environments.

330. The PIREP programme has been developed to implement an extended version of the optimal control model (Doyle, Hoffman, 1976). As opposed to the conventional model, PIREP incorporates certain effects of visual instrument scanning and attention allocation (including an optimising procedure) as well as motion cues sensed by the vestibular system. As opposed to programme MANMOD, PIREP is restricted to the study of time-independent effects and is not designed for interactive operation.

331. The optimal control model has been implemented with a general purpose control systems analysis programme (Program CC: Thompson, 1985) that runs on IBM-PC compatible computers. Program CC includes a comprehensive selection of tools and algorithms for frequency domain analysis, time domain analysis, sampled data systems, multi-input/multi-output systems, state space methods, and optimal control procedures. Program CC provides two human performance modelling approaches, (1) the optimal control model and (2) the quasilinear models and McRuers's law described in section 3.3.9 (see also Allen et al., 1989).

332. The OCM has been implemented as the SSOCM (Steady-State Optimal Control Model) computer programme for operation on a Digital Equipment Corporation VAX machine using the VMS operating system and for IBM PC-, XT-, and AT-compatible personal computers using the DOS environment (Levison, 1989). The SSOCM software system is used to predict operator/vehicle behaviour in linear, steady-state control tasks. A model for task interference and attentional workload is incorporated in the programme, and perceptual limitations such as thresholds and resolution limitations can be accounted for. This implementation of the "steady-state" model treats operator/vehicle tasks in which all problem variables may be considered as zero-mean Gaussian processes having stationary statistics. The steady-state model implementation takes advantage of the mathematical properties of linear systems driven by Gaussian noise to yield directly the statistics of the problem solution. The problem solution is, of course, consistent with the conceptual model described above. The inputs to the steady-state model consist entirely of parameters that describe the task environment and the operator, as described in the preceding discussion of the conceptual model. Because no time histories are generated, there are no input signals directly analogous to the external forcing-function time histories that would be required in a simulation experiment. In the steady-state model implementation, the operator's internal model of the task environment must be identical to the linear model of the task environment.

References

Allen, R. W., McRuer, D. T., & Thompson, P. M. (1989). Dynamic systems analysis programmes with classical and optimal control applications of human performance models. McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Baron, S. & Kleinman, D. L. (1969). The human as an optimal controller and information processor. IEEE Trans. Man-Machine Systems, 10, (1), 9-17.

Baron, S. & Berliner, J. E. (1974). MANMOD: A computer programme for statistical analysis of dynamical systems involving man as controller. (BBN Report 2704). Cambridge, MA: BBN Laboratories.

Baron, S. & Levison, W. H. (1975). An optimal control methodology for analysing the effects of display parameters on performance and workload in manual flight control. IEEE Trans. Syst., Man, and Cybern. 5, (4), 423-430.

Baron, S. & Levison, W. H. (1977). Display analysis with the optimal control model of the human operator. Human Factors, 19, (5), 437-457.

Baron, S. & Levison, W. H. (1980). The optimal control model: Status and future directions. Proceedings, IEEE Conference on Cybernetics and Society. New York: IEEE

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In W. B. Rouse (Ed.), Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Boff, K. R. & Lincoln, J. E. (Eds.) (1988). Engineering data compendium, human perception and performance (4 Volumes). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Doyle, K. M. & Hoffman, W. C. (1976). Pilot modelling for manned simulation: Program PIREP user manual. (AFFDL-TR-76-124). Wright-Patterson Air Force Base, OH: Flight Dynamics Laboratory.

Gai, E.G. & Curry, R.E. (1976). A model of the human observer in failure detection tasks. IEEE Trans. on Syst., Man, and Cybern., 6, (2), 85-94.

Hess, R. A. (1977). Prediction of pilot opinion ratings using an optimal pilot model. Human Factors, 19, (5), 459-475.

Hess, R. A. (1987). Feedback control models. In: G. Salvendy (Ed.), Handbook of human factors. New York: J. Wiley.

Kleinman, D. L., Baron, S., & Levison, W. H. (1971). A control theoretic approach to manned-vehicle systems analysis. IEEE Trans. Autom. Control, 16, 824-832.

Kleinman, D.L. & Curry, R.E. (1977). Some new control theoretic models of human operator display monitoring. IEEE Trans. Syst., Man, and Cybern., 7, (11), 778-784.

Korn, J. & Kleinman, D. L. (1978). Modelling the effects of high-G stress on pilots in a tracking task. Proc., Fourteenth Annual Conference on Manual Control. (NASA Conf. Pub. 2060). Moffett Field CA: NASA/Ames Res. Cent.

Levison, W.H. & Tanner, R. B. (1971). A control-theory model for human decision making. Seventh Annual Conf. Manual Control, (NASA SP-281). Washington, DC: National Aeronautics and Space Administration.

Levison, W. H. & Houck, P. D. (1975). Guide for the design of control sticks in vibration environments. (AMRL-TR-74-127). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory.

Levison, W. N. & Junker, A. M. (1977). A model for the pilot's use of motion cues in roll-axis tracking task. (AMRL-TR-77-40). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratory.

Levison, W. H. (1978). Model for human controller performance in vibration environments. Aviation, Space, and Environmental Medicine, 49, (1), 321-327.

Levison, W. H. & Warren, R. (1984). Use of linear perspective cues in a simulated height regulation task. Proc., Twentieth Annual Conference on Manual Control. (NASA Conf. Pub 2341). Moffett Field CA: NASA/Ames Res. Cent.

Levison, W. H. (1989). The optimal control model for manually controlled systems. McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

McRuer, D. T. (1980). Human dynamics in man-machine systems. Automatica, 16, 237-253.

Moray, N. (1986). Monitoring behaviour and supervisory control. In: K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Pew, R.W., Baron, S., Feehrer, C.E., & Miller, D.C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation. (BBN Report 3446). Cambridge, MA: BBN Laboratories.

Phatak, A.V. & Kleinman, D.L. (1972). Current status of models for the human operator as a controller and decision maker in manned aerospace systems. Proceedings, Automation in Manned Aerospace Systems. (CP-114). Neuilly-sur-Seine, France: AGARD.

Schmidt, D. K. (1979). Optimal flight control synthesis via pilot modelling. J. Guidance and Control, 2, 308-312.

Stein, W. (1989). Models of human monitoring and decision making in vehicle and process control. McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Stengel, R. F. & Broussard, J. R. (1978). Prediction of pilot-aircraft stability boundaries and performance contours. IEEE Trans. Syst., Man, and Cybern., 8, 349-356.

Thompson, P. M. (1985). Program CC version 3 user's guide. Hawthorne, CA: Systems Technology, Inc.

Future Needs

333. Although developed 20 years ago, the application of the optimal control model still is in its early stages and many questions remain to be answered. There is need for additional data and appropriate experimental paradigms. The major opportunity for extending the model to other tasks is in the area of supervisory control. Although highly developed as a useful analytical tool, further model development is recommended to both enhance predictive capabilities for current applications and to extend the OCM to new applications and new kinds of tasks. Three areas for future research are proposed: (1) improve predictive accuracy for high-order continuous control tasks, (2) develop models for skill acquisition and (3) extend to additional tasks.

334. As noted in the discussion on model application, reliable performance trends can be often obtained using a fixed set of values for pilot-related model parameters. Detailed measurements of pilot response behaviour in a variety of laboratory tasks have been obtained using these values. Nevertheless, if we consider the full range of laboratory results, certain systematic deviations in these parameters are observed. Specifically, motor time constants greater than 0.1 seconds have been observed in tasks requiring the control of plants that include high-order dynamics or significant delays. In addition, noise/signal ratios well below the nominal value of -20 dB have been observed in certain situations, involving mainly control of unstable dynamics. Additional study is also desired to enhance the predictive accuracy for tasks involving transient inputs. The OCM has a structure that is suitable for treating deterministic inputs; but, as explained earlier, predictions that one obtains with the model are highly dependent on the assumptions made concerning the operator's knowledge of the input.

335. Development of models for learning behaviour (i. e., control-strategy development) is one of the important remaining theoretical frontiers in manual control. Models of this sort have ready application to the design of training simulators. Because this area of research (at least with regard to OCM application) is in its early stage, it is not clear what directions it will take. One approach is simply to construct an outer loop about the OCM as currently structured; for example, discover a suitable set of rules for adjusting the pilot-related model parameters from a set of values appropriate to some initial state of training to another set of values typical of fully-trained pilots. A more analytically-oriented approach would be to include a fourth adaptive element to the pilot model: an optimal identifier to account for development of the pilot's internal model in a given task situation. This model element would account both for the rate of learning as well as the asymptotic structure of the internal model.

336. The major opportunity for extending the OCM to other tasks is in the area of supervisory control (see the models DEMON and PROCRU in Chapter 4 of this report). These control problems, as we have noted, involve monitoring, detection, decision-making, and discrete and/or infrequent control. Most often, the systems are highly automated, require more than one operator and are extremely complex. The principal feature of the OCM that is useful for these applications is its information processing structure, but the underlying, normative modelling perspective is also important. The application of the OCM, or its derivatives to such problems, is in its early stages and many questions remain to be answered. Among the more important are questions concerning the human's internal model for such large scale systems, the appropriate control and decision cost functionals for these problems, the modelling of attention-sharing strategies in time-varying situations, and the appropriate level for incorporating and modelling aspects of the tasks that are important but are not likely to be treated using the same techniques. It must be recognised that validation of such complex models to the degree that manual control models have been and can be validated, will not be possible for some time, for both theoretical and practical reasons. One may not expect, therefore, that supervisory control models that are predictive, in the same sense as the OCM, will be developed in the near future. Nonetheless, it should be possible to develop models of supervisory control that will capture many essential features of tasks of interest and will prove to be useful design, analysis, and evaluation tools.

3.3.11 A Model of the Helicopter Pilot

Summary Description

337. The pilot model was developed to control a helicopter whose system dynamics were fourth order. The decision process of the pilot is represented by a hierarchical network. Implicit in the design of the model is that the pilot samples the control variables periodically. On receiving the sampled-data variables, the hierarchical decision network determines the appropriate multi-loop closure and tracking characteristics for the whole system. It is assumed that the pilot stabilises each variable in order with loop closure from inside out, the order being roll angle, velocity, and lateral deviation (position). Depending on the stability level of the system, the tracking characteristics can be high bang-bang, medium bang-bang, or simple tracking (gain plus lead). The tracking characteristics are implemented through a neuromuscular dynamic model.

338. The model of a helicopter pilot was studied using simulation by Benjamin (1970). The basic element of the pilot model is a decision hierarchy which determines the multiloop closure and tracking characteristics of the man-vehicle system. Pilot model input is quantised and used by the hierarchy to determine the specific loop to be closed and the particular transfer function to apply to that loop. The pilot model and vehicle dynamics are implemented on a computer. Model validation is provided by comparison of tracking records obtained from this simulation of the vehicle with a human operator. Although developed for a vehicle with only two lateral degrees of freedom, the pilot model is sufficiently general in form to allow its extension to six degrees of freedom. As a fourth-order system, it is applicable to the control of not only the helicopter, but all VTOL vehicles. The reduction of higher-order inputs to zero permits applicability to vehicles with lower-order dynamics. Its form is independent of the input function.

History and Source

339. The model was developed by P. Benjamin (1970) while at Northwestern University, Evanston, Illinois, USA.

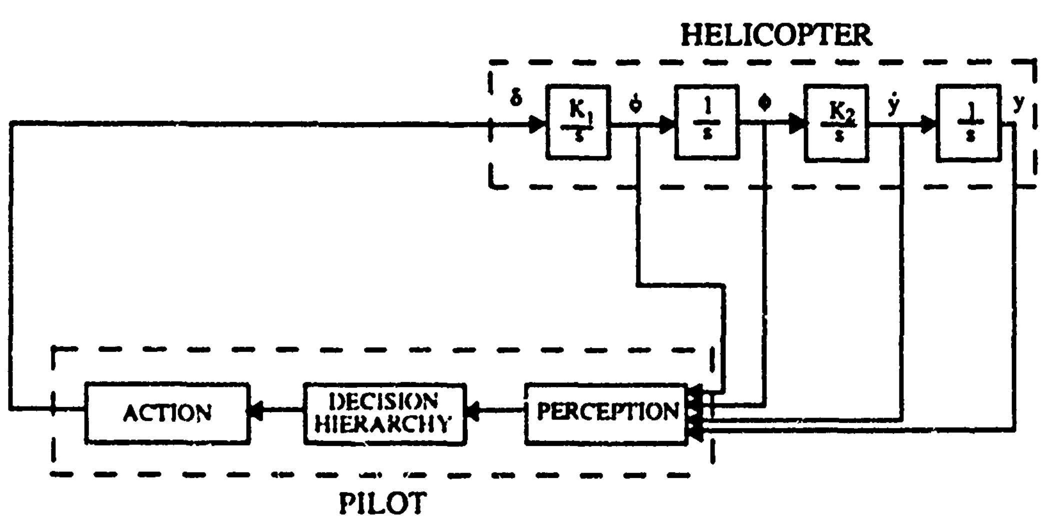


Fig. 3.16 The helicopter pilot model

Product and Purpose

340. The results from the pilot model correlated reasonably well with those obtained from a fixed base simulator. Therefore the model can be assumed to duplicate the control action of a helicopter pilot attempting to maintain a certain lateral position.

341. As shown in Figure 3.16, the angular accelerations of a helicopter are governed by the control stick. Without compensation in any form, a change in stick position, d, causes a change in the angle of attack of the rotor blades. This, in turn, results in a moment about the centre of mass of the vehicle, the second integral of which is the roll angle, f. Since the lateral thrust of the helicopter is proportional to the roll angle, the relation between the angle of attack of the vehicle and an inertial position, y, also involves two integrations. Thus, the relationship between control stick deflection and lateral position is a fourth-order function.

342. The pilot of a helicopter observing the world outside of his vehicle, or, in the case of this experiment, the subject viewing his visual display, receives information with respect to the current values of roll rate, roll angle, velocity, and position of his aircraft. As diagrammed in Figure 3.17, he must evaluate this input information, perform some decision function based upon the results of this evaluation, and then effect some appropriate movement of the control stick. This pilot model considers the first of these, evaluation, in terms of sampling and quantification of the input variables. The decision function is modelled as a hierarchical network, and action is taken in terms of a neuromuscular dynamics model. In this pilot model, roll, velocity, and position are quantified in four levels by three threshold values. The "panic" threshold separates "very high" from "high" levels. The "maximum" threshold divides "high" and "acceptable" levels, and the "minimum" threshold draws the line between "acceptable" and "negligible" values of the three variables. Roll rate is not quantified.

343. To control a fourth-order system, such as the helicopter used in this experiment, the human operator must generate a fourth-order lead. It has been shown, however, that operators are generally capable of producing second-order or sometimes third-order lead. It appears that the successful human controller of a higher-order system utilises a hierarchy of control techniques. That is, although keeping track of the values of all variables, he tends to stabilise each variable in order, with loop closure proceeding from the inside out. The first order of the hierarchy is to maintain low roll angles. Once the angular variation has been stabilised, the subject attempts to reduce his velocity to near zero. Once he has stabilised himself at some position away from the required hover point, he attempts to correct his lateral deviation. Of course, while stabilising roll he cannot allow a large positional deviation, but once each order has been stabilised, maintaining stability becomes more of a simple gain task than a lead task. Thus the complexity of the control task has been reduced by a simple expedient of stabilising each loop successively from the inside out.

344. The full decision hierarchy is summarised in flow diagram form in Figure 3.17. Note that only one type of control is effected (only one loop closed at a time) corresponding to the highest level of the hierarchy which has input values in the range indicated. The result of the decision hierarchy is a decision as to which loop to close and what type of tracking to use. This defines a desired stick position, or a goal, toward which the stick is moved. Young and Meiry (1965) have found that in control of higher-order systems, the less stable the system the more the human controller utilises a bang-bang type of control. This refers to the type of control in which the operator moves the stick in discrete jumps and maintains a constant stick position between jumps. In a previous study of helicopter control, the author noted that in less stable situations requiring first- or second-order leads on the part of the operator, bang-

bang control was elicited. As successive loops became stabilised and the requirement switched increasingly toward a gain, the control stick movements tended away from bang-bang control toward simple tracking.

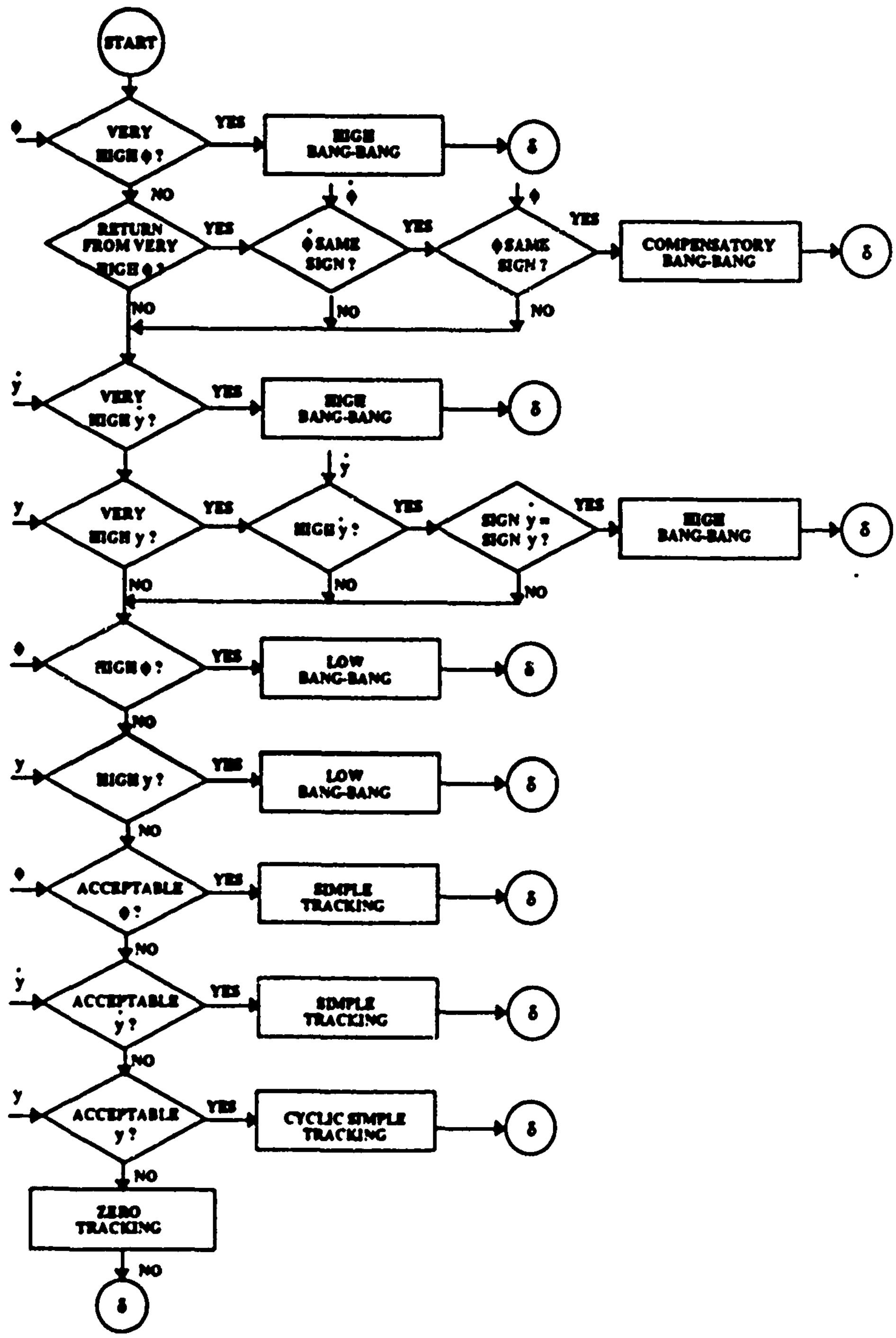


Fig. 3.17 Decision flow diagram.

345. The proposed helicopter pilot model uses dual-level bang-bang tracking and simple tracking (gain plus lead) for each loop, depending upon the stability level of the closed loop. The hierarchical controller, then, interprets the quantified input, decides which loop is to be closed, and determines the mode of tracking to be unutilised in the control of this closed loop. Thus, if roll angle is in the very high range, a decision is made to use high-level bang-bang tracking. For high roll angle errors, low-level bang-bang tracking is utilised, and in the acceptable range of roll, simple tracking suffices. For negligible level roll angles, the roll loop is open.

346. The use of a decision hierarchy as a determinant of the multiloop closure and tracking characteristics represents the major contribution of this model. Thus the concept of the decision hierarchy and the stabilisation of the multiloop system by successive single-loop closures constitute the primary contributions of this model.

When Used

347. To be used at the experimental design stage.

Procedures for Use

348. Not available.

Advantages

349. By using a digital simulation approach a complex man-machine system can be analysed in a relatively straightforward manner. The form of the hierarchical decision network is independent of the input function and can therefore be applied to other systems.

Limitations

350. This type of model requires considerable computer facilities for synthesis and analysis.

Application Examples

351. Benjamin (1970) used the model to control the lateral position of simulated helicopter high-order dynamics. It is claimed the structure of the model can easily be adapted for controlling all VTOL Vehicles.

Technical Details

352. Not available.

References

Benjamin, P. (1970). A hierarchical model of a helicopter pilot. Human Factors, 12, (4) 361-374.

Young, L. R. & Meiry, J. L. (1965). Bang-bang aspects of manual control in higher order systems. IEEE Trans. on Automatic Control, 10, 336 - 341.

Future Needs

353. Although the pilot model was developed in a somewhat restricted context, the system which was investigated and the final model are general enough in form to allow extensions, e. g., to a full six degrees of freedom.

3.3.12 The Veldhuyzen Helmsman Model

Summary Description

354. The helmsman model consists of an internal model plus decision-making (logic) elements. The internal model is based on a psychological concept (Cooke 1965) which postulates that a human operator must have knowledge about the dynamic behaviour of a system in order to be able to control it. A survey of the internal model principle is given by Wonham (1976). In this case the internal model is represented by a linear differential equation and is used to predict the response of a ship to the helmsman's actions. The logic element then decides whether the action is having the desired effect or not. It also decides when and how much action is required according to certain limiting parameters. Thus the model is algorithmic in nature. In estimating the parameters for the model a parameter tracking method was used.

History and Source

355. The model was developed by Veldhuyzen (1976). The work arose out of a need for validated helmsman models.

Product and Purpose

356. The purpose of the model is to provide ship designers with a tool for examining the closed loop responses and handling qualities of a proposed design. For the time domain results presented, the model reproduced control action similar to that of a helmsman. Also the heading of the ship steered by the model closely matched the ship's heading steered by a helmsman.

357. In order to understand the nonlinear helmsman's model, some remarks about the task the helmsman has to perform should be made. In general, the helmsman's task may be considered to be a pursuit tracking task, where the input signal (the headings ordered) consists of a series of steps of randomly distributed amplitudes and durations. The model is based directly on the internal model concept and consists of the following parts:

(1) an internal model equation to make it possible to predict future ship states,

(2) a decision making element in order to generate the helmsman's actions in controlling the ship including the use of the predictions of the internal model,

(3) an estimator to estimate the state of the ship from the displayed heading disturbed with noise due to waves.

358. In most of the experiments considered the helmsman's decision-making process during the manoeuvre can be divided into four phases. During the first phase, the helmsman generates an output in order to start the ship rotating , then during the second phase, the rudder is kept centred. During the third phase, the helmsman stops the rotating motion of the ship and when the desired heading is achieved with only a small rate of turn (the desired or ordered state), the fourth phase starts, with a rudder angle of zero. If the rate of turn is not small, there will be an overshoot, and then to achieve the desired state, the cycle has to be

repeated, starting with the first phase again. The four phases can be indicated in the phase-plane, that is, the plot of the rate of turn as a function of the heading error.

359. After the rudder angle is chosen, the internal model is used to determine whether the objectives will be satisfied during the following period, or whether a new rudder angle has to be chosen. For large heading errors, the helmsman is mainly interested in reducing the instantaneous error; whereas for small errors, he is particularly focused on the difference between the predicted state of the ship and the objectives to be reached. With a few exceptions, the experimental values of the performance measures can be reasonably predicted by the computer simulations.

When Used

360. At the conception stage of a ship's design.

Procedures for Use

361. To predict the performance measures by means of computer simulations with the nonlinear model, the structure of the model must be adapted to the displays used during a particular test. The information supplied by the rate of turn indicator can be used as an initial condition to make predictions with the internal model.

Advantages

362. This model is one of the few validated helmsman models available.

Limitations

363. The model is limited to describing the course-keeping control characteristics of a helmsman.

Application Examples

364. The model was used successfully to control the simulated dynamics of a supertanker during course-keeping (Veldhuyzen, 1976, 1977). It was also used to model the control characteristics of a helmsman during full scale trials on board a small ship. Under the later circumstances the model did not perform particularly well. The nonlinear helmsman's model has been used in a number of situations to analyse the helmsman's behaviour (Veldhuyzen, 1976). The model has been used to study the influence of additional information presentation systems on the performance of the helmsman steering a directionally stable or unstable ship by means of computer simulations. The influence of this auxiliary equipment on the helmsman's performance in relation to the dynamics of ships has also been investigated. With each ship, the following manoeuvres were simulated:

(1) Manoeuvre 1: Course keeping.

(2) Manoeuvre 2: The execution of a heading order of 5 degrees

(3) Manoeuvre 3: The execution of a heading order of 25 degrees.

Technical Details

365. Not available.

References

Ackermann, D. & Tauber, M. J. (Eds.). (1990). Mental models and human-computer interaction 1. Amsterdam: North-Holland.

Cooke, J.E. (1965). Human decision in the control of a slow response system. D. Ph. Thesis, Oxford University, UK.

Gentner, D. & Stevens, A. L. (Eds.). (1983). Mental models. Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N. (1983). Mental models. Cambridge, UK: Cambridge University Press.

Kelley, C. R. (1968). Manual and automatic control. New York: J. Wiley.

Rasmussen, J. (1990). Mental models and the control of action in complex environments. In D. Ackermann & M. J. Tauber (Eds.). Mental models and human-computer interaction 1. Amsterdam: North-Holland.

Stassen, H. G., Johannsen, G., & Moray, N. (1990). Internal representation, internal model, human performance model and mental workload. Automatica, 26, (4) 811-820.

Veldhuyzen, W. (1976). Ship manoeuvring under human control. Ph.D. Thesis, Delft University of Technology, The Netherlands.

Veldhuyzen, W. & Stassen, H. G. (1976). The internal model: What does it mean in human control?In T. B. Sheridan & G. Johannsen (Eds.). Monitoring behaviour and supervisory control. New York: Plenum Press.

Veldhuyzen, W. & Stassen, H. G. (1977). The internal model concept: An application to modelling human control of large ships. Human Factors, 19, (4) 367-380.

Wilson, J. R. & Rutherford, A. (1989). Mental models: Theory and application in human factors. Human Factors, 31 (6), 617-634.

Wonham, W.M. (1976). Towards an abstract internal model principle. IEEE Trans. Syst., Man, and Cybern., 6, (11) 735-740.

Future Needs

366. This example demonstrates how the internal model concept can contribute in building models to understand the behaviour of the human operator and in the analyses of additional displays. Many important problems in human operator activities can be directly related to the internal model concept, e. g., mental load problems and the problems involved in monitoring and decision making, especially in the context of slowly responding systems.

3.3.13 A Model of Visual Scene Perception in Manual Control

Summary Description

367. A model of visual scene perception in manual piloting of an aircraft was developed by Wewerinke (1980). After theoretical and experimental studies and a concise inventory of the most important visual scene characteristics, the visual perception process is modelled on the basis of the linear perspective geometry and cues related to the relative movement of the observer with respect to the outside world. This involves mathematical relationships between visual cues and vehicle state variables. After linearisation, the model can be integrated into the existing framework of the optimal control model (OCM) describing manual control behaviour. The visual scene perception model involves assumptions concerning perceptual thresholds of the various cues, noise levels associated with observing these cues, and the interference among them. Values for these parameters were derived from baseline experiments supplemented with data from the psychophysical literature.

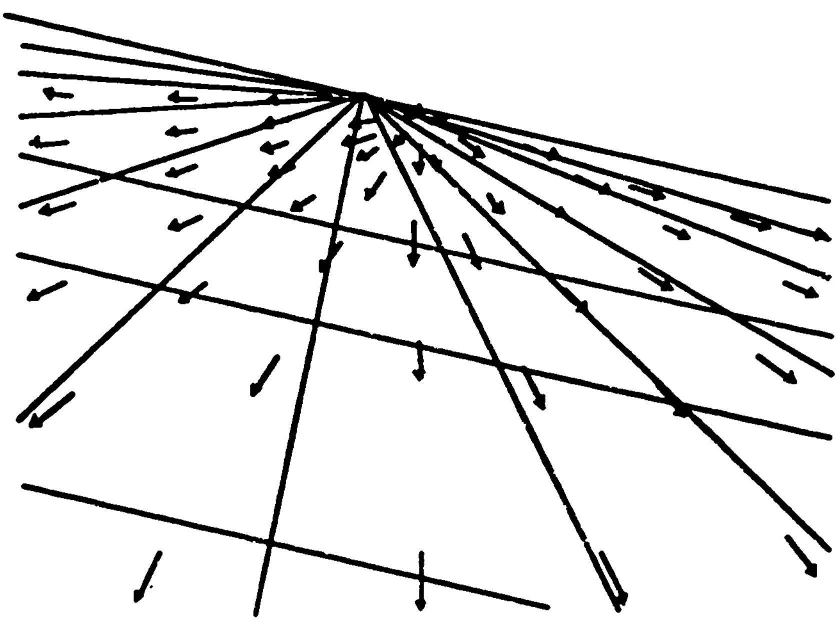


Fig. 3.18 Visual scene

History and Source

368. Although the literature on visual perception is vast, a selective view seems adequate to arrive at the most important visual cues involved in the visual scene perception process. One of the earliest studies on visual scene perception directly related to flight control problems was performed by Gibson (1950). According to Gibson, the most important visual cues which can be derived from the visual scene are related to:

(1) linear perspective geometry

(2) relative motion or motion parallax

(3) apparent size of objects whose real size is known

(4) occultation of a far object covered by a near one

(5) distribution of light and shade over an object

(6) aerial perspective and the loss of detail with distance.

369. It is commonly assumed that linear perspective geometry and relative motion provide important cues to distance and motion perception. A schematic version of the visual scene is shown in Figure 3.18 consisting of lines and points (textural elements). The point of the visual field toward which the observer is moving (rectilinearly) appears to be stationary. All other textural points move with respect to the observer and can be indicated by velocity vectors ("streamers").

370. The perspective interpretation of the elements of the visual scene allows an estimation of the linear and angular position and velocity of the observer. This involves not only the momentary information provided by the visual scene cues but also past information. The resulting dynamic and stochastic process can be described in estimation theoretical terms. It involves a mathematical description of the visual cues and their functional relationship as well as expectations of the human observer.

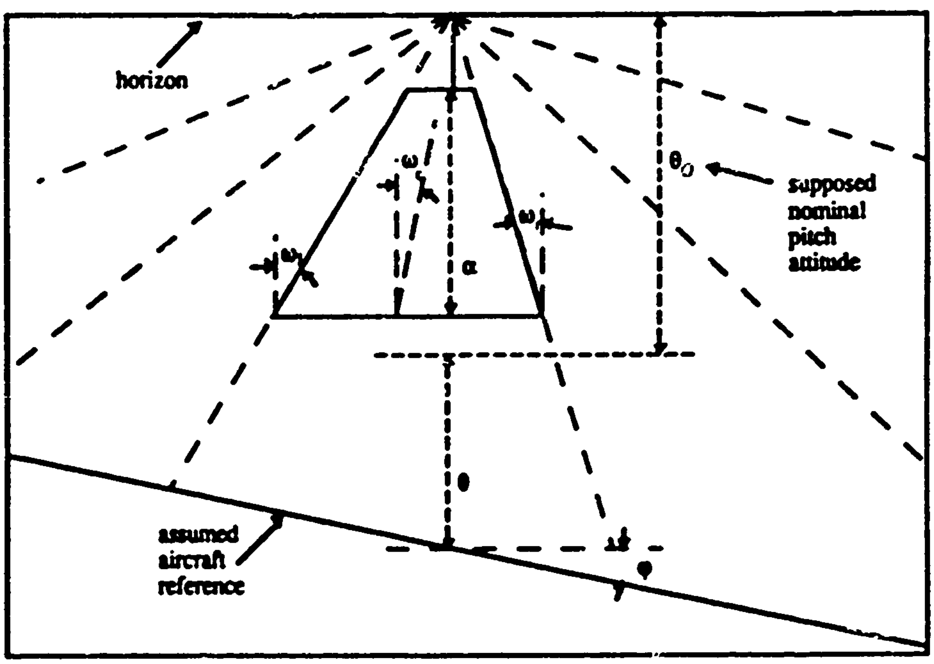


Fig. 3.19 Cues derived from the visual approach scene.

Product and Purpose

371. The scene perception model is an extension of the optimal control model (OCM) presented in section 3.3.10. It is based on the fundamental hypothesis that the well-motivated, well-trained human operator behaves in a near optimal manner subject to his inherent constraints. This implies that the description of human behaviour is concentrated on

two aspects: subjective criteria for optimality and the human limitations. In addition, it is assumed that the human operator is dealing with a linear (-ized) system. Once these assumptions are made, linear optimization and estimation theory can be used to formulate the various aspects and stages of human information processing.

372. A schematic version of the visual approach scene, relevant for the modelling approach, is shown in Figure 3.19. The cues which are assumed to be provided by this scene are indicated. The most important cue for lateral guidance is derived from the inclination of the runway sides and/or the runway centreline. The lateral deviation is zero if the inclination of both runway sides is the same ($w_r = w_l$) and the inclination of the centreline is zero ($w_c = 0$). Vertical guidance must be based on the average inclination of the runway sides when no runway end and no horizon is visible. In that case, the observer must know the nominal inclination (which is range-varying). However, a better indication of the vertical position can be obtained when the projected length of the runway a (or, almost equivalently, the depression of the runway threshold with respect to the horizon) is visible.

373. The first aspect in the model development procedure concerns the task environment. Next, it is described how the human operator perceives this environment and processes the perceived information resulting in an internal representation of this environment. Finally, it is briefly indicated how the human operator involved in a manual control task utilises this internal representation for his control response behaviour. Further details of the optimal control model can be found in the appropriate summary description (section 3.3.10).

When Used

374. For the design or evaluation of external scenes used in flight simulation.

Procedures for Use

375. Not available.

Advantages

376. The visual scene provides a variety of perspective geometrical and relative motion cues. Experimental results support the hypothesis that these characteristics can be considered as separate cues among which the human operator must divide his attention. Both the workload model results and the subjective ratings indicate that human operator workload is indeed increased when performing control tasks simultaneously.

377. The visual scene perception model has been shown capable of providing a general framework for dealing quantitatively with the important visual scene characteristics. Integrating this model with the optimal control model (OCM) of human control behaviour allowed model predictions to be made using a priori perceptual threshold values for the various visual cues involved. The experimental results in terms of mean-squared system output scores agreed relatively well with those of the model, and showed the predictive capability of the model. The more detailed frequency domain measures (human describing functions and observation noise spectra) allowed for refined estimates of the perceptual thresholds.

Limitations

378. Limitations given for the OCM apply here also.

Application Examples

379. To validate the visual scene perception model an experimental programme was conducted. A variety of visual approach scene conditions were involved to provide a critical test for the hypotheses and assumptions underlying the model results. The reader is referred to Wewerinke (1978, 1980) for more detailed information. Ten task configurations were selected representing various visual scene conditions and control modes. The vertical and/or lateral rate control tasks were stationary, i.e., it was assumed that the aircraft was "frozen" at a fixed point along the approach path at a nominal altitude of 200 ft for a 3° glideslope. In this way, detailed information concerning the pilot's information processing associated with the various visual scene conditions could be obtained from mean-squared system variables and frequency domain measures such as human describing functions and observation noise spectra. Especially the latter provided a sensitive check on the exactness of the values used for the model parameters under investigation.

380. The experimental results agreed well with the model predictions. For all single-axis tasks the experimental scores lay well within the predicted interval. This suggests not only that the model is "right" but also that the assumed numerical values for the thresholds and overall attention are close to the "real" values. For the dual-axis tasks the model predictions were somewhat pessimistic. An additional adjustment of the originally assumed model parameters was necessary to obtain a better model match to the attitude and control scores.

381. Several pilot workload model predictions are supported experimentally showing the usefulness of the model to analyse the cost of performing the manual approach task. This is essential, as pilot workload is often the most sensitive variable. Also, the workload model predictions have been confirmed by subjective ratings. Apart from one configuration, the linear correlation between model predictions and subjective ratings was 0.8.

Technical Details

382. Model parameters can be divided into parameters which are constant for all configurations and parameters which were considered as the remaining model variables. The key variables were the perceptual thresholds. As discussed before, it was assumed that the human operator divides his attention optimally, i.e., minimising the given cost functional, among the visual cues. The other model parameters were fixed across configurations and chosen on the basis of previous studies: a perceptual time delay of 0.2 s, a neuromotor time constant of 0.1 s and a motor noise ratio of -30 dB; overall level of attention P_0 was set at -20 dB, and varied plus/minus 2 dB to determine performance sensitivity.

References

Baron, S. & Levison, W. H. (1980). The optimal control model: Status and future issues. Proceedings, IEEE Conf. on Syst., Man, and Cybern. New York: The Institute of Electrical and Electronics Engineers.

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In W. B. Rouse (Ed.). Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Baron, S. (1987). Application of the optimal control model to assessment of simulator effectiveness. In W. B. Rouse (Ed.). Advances in man-machine systems research (Vol. 3). Greenwich, CT: JAI Press.

Gibson, J. J. (1950). The perception of the visual world. Cambridge, MA: The Riverside Press.

Grunwald, A. J. & Merhav, S. J. (1976): Vehicular control by visual field cues - Analytical model and experimental validation. IEEE Trans. on Systems, Man, and Cybernetics,. 6, (12) 835-845.

Haber, R. N. (Ed.) (1968). Contemporary theory and research in visual perception. New York: Holt, Rinehart, and Winston.

Wewerinke, P. H. (1978). Visual scene perception process involved in the manual approach (NLR TR 78 130 U). Amsterdam: National Aerospace Laboratory NLR.

Wewerinke, P. H. (1980): Visual scene perception in manual control. Journal of Cybernetics and Information Science. 3, 3-26.

Future Needs

383. It is expected that the present knowledge about the visual perception process provides a useful guideline for modelling the perception of new visual scene situations. This offers the possibility for dealing quantitatively with the effect of, for instance, visibility conditions, runway or road characteristics and vehicle reference information. The integrated modelling approach also provides the capability of investigating the interaction of these characteristics with other display information, vehicle characteristics, disturbance environment, etc.

3.3.14 A Fuzzy Set Model of the Car Driver

Summary Description

384. Fuzzy set theory has been applied to the development of a model of car driver behaviour during lane keeping manoeuvres. It was considered that a driver is always uncertain in the recognition of the external environment of the car and in the selection of the most appropriate controller action. The uncertainty of the driver is generated by a lack of perfect conformity between his own internal model and any external stimuli. A basic postulate of the car driver model is that a human is capable of parallel processing of visual information. The visual system of a car driver is assumed to be capable of separating a visual pattern according to its spatial frequency components. This form of pattern recognition is used to classify simulated road patterns as more or less distinct right or left curves. A fuzzy classification is made of the distinctiveness of the curves which in turn is used to generate fuzzy motor commands to the steering wheel via a dynamic reaction model.

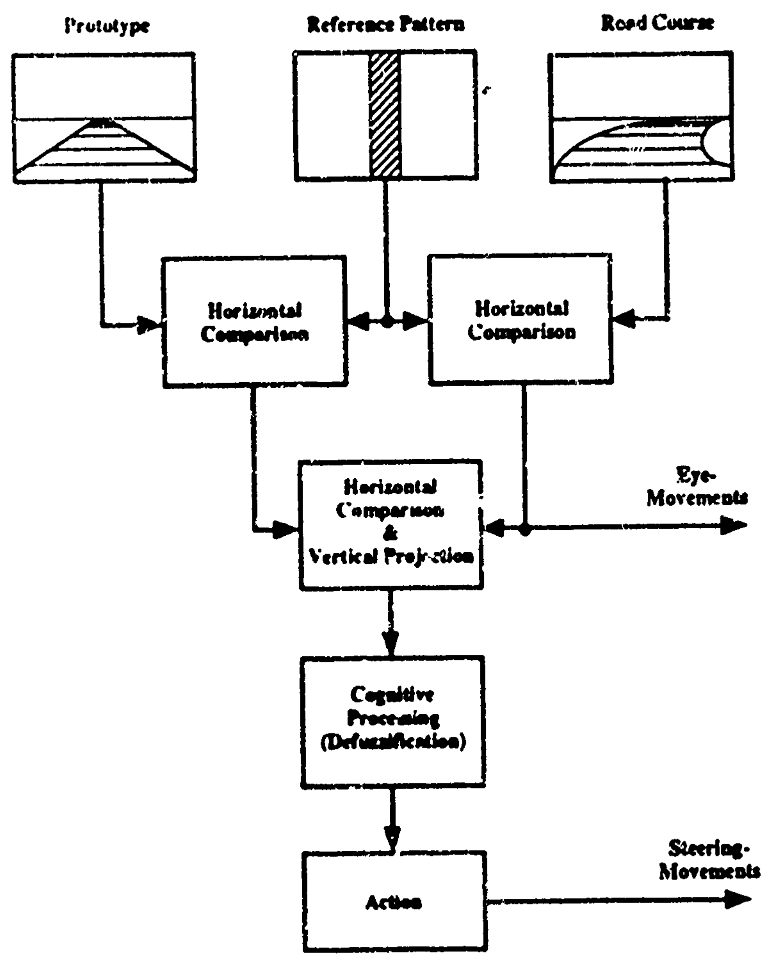


Fig. 3.20 Fuzzy model of driver behaviour in lane-keeping (Willumeit, Kramer and Rohr, 1983).

History and Source

385. The model was developed by Kramer (1985) at the Technical University of Berlin, FRG.

Product and Purpose

386. The aim of the model is to provide a simulation facility which describes the driver-vehicle environment in order to be able to evaluate the dynamics of a vehicle without full scale trial runs with an experimental vehicle for each variation of dynamical parameters. The simulation programme has two operating modes, i.e., open-loop mode and closed-loop mode. The open-loop mode reconstructs the visual field of the driver in accordance with the data for the relative position of the vehicle on a given road course in terms of horizontal eye-movements. The closed-loop mode generates a time history of the defuzzified steering movement which acts as the input to the car dynamics.

387. Steering movements are expected to have an adequately close relationship with eye-movements; but it must also be taken into account that the steering behaviour is influenced more strongly by internal models and concepts of the driver than the fixation behaviour. Figure 3.20 shows the driver model being composed of parallel branches of processing stages such that steering movements do not only follow from the horizontal component of eye-movements, but also from parallel processing of additional, internally stored instructions. For the particular task of lane keeping, the instruction refers to the permitted lane which has to be chosen as well as to the admissible lane deviation. The corresponding pattern is then selected in such a way that if it coincides with the actual pattern of the road, no steering reaction is affected. It is, therefore, called a prototypical pattern.

When Used

388. At the design stage in the development of a prototype vehicle.

Procedures for Use

389. Not available.

Advantages

390. The fuzzy set theory approach makes the model flexible with regard to increasing its sophistication.

Limitations

391. The model needs to be extended in order to model avoidance of lane obstructions.

Application Examples

392. The car driver model has been validated against simulator results. The results showing adequate conformity between the model and the simulation.

Technical Details

393. Not available.

References

Kramer, U., Rohr, G., (1982). A model of driver behaviour. Ergonomics, 25, (10) 891-907.

Kramer, U. (1985). On the application of fuzzy sets to the analysis of the system driver-vehicle environment. Automatica, 21, (1) 101-107

Future Needs

394. Not available.

3.3.15 Fuzzy Set Models of a Helmsman

Summary Description

395. Models that describe the control characteristics of a helmsman engaged in course-keeping and course-changing tasks have been developed. The basic element in each of the models is a fuzzy rule-based algorithm. The models are used in digital computer simulations to control the nonlinear yaw dynamics of a Royal Navy warship model which is subjected to sea state disturbances. In both models, fuzzy sets on

(1) yaw error ε

(2) yaw rate $\dot{\psi}$, and

(3) rudder demand δ

were used. The fuzzy sets used to describe the course-changing characteristics were taken as:

(1) NB = negative big,

(2) NS = negative small,

(3) AZ = approximately zero,

(4) PS = positive small, and

(5) PB = positive big.

396. Thus the fuzzy course-changing algorithm describing the helmsman's action is assumed to be:

1. If ε is PB and $\dot{\psi}$ is any	then δ is PB else
2. If ε is PS and $\dot{\psi}$ is NS or NB	then δ is PB else
3. If ε is PS and $\dot{\psi}$ is PS or AZ	then δ is PS else
4. If ε is PS and $\dot{\psi}$ is PB	then δ is AZ else
5. If ε is AZ and $\dot{\psi}$ is NB	then δ is PB else
6. If ε is AZ and $\dot{\psi}$ is NS	then δ is PS else
7. If ε is AZ and $\dot{\psi}$ is PS	then δ is NS else
8. If ε is AZ and $\dot{\psi}$ is PB	then δ is NB else
9. If ε is NS and $\dot{\psi}$ is NB	then δ is AZ else
10. If ε is NS and $\dot{\psi}$ is NS or AZ	then δ is NS else
11. If ε is NS and $\dot{\psi}$ is PS or PB	then δ is NB else
12. If ε is NB and $\dot{\psi}$ is any	then δ is NB.

397. The output fuzzy set is then defined by:

$$\mu(\delta_c) = \max_{n=1\ldots 12}\left[\min\left\{\mu_{\varepsilon_n}(\varepsilon), \max\left(\mu_{\dot\psi_{n_1}}(\dot\psi), \mu_{\dot\psi_{n_2}}(\dot\psi)\right), \mu_{\delta_n}(\delta)\right\}\right] \tag{1}$$

398. In order that the helmsman's control output is a deterministic value, the centre of area procedure is applied to the output fuzzy set.

399. In the course-keeping algorithm, the fuzzy sets on yaw error e and rudder demand d were taken as:

(1) PVS = positive very small,

(2) PVVS = positive very very small,

(3) Z = zero,

(4) NVS = negative very small, and

(5) NVVS = negative very very small.

400. The yaw rate $\dot\psi$ being defined as:

(1) FI = increasing fast,

(2) IS = increasing slow,

(3) Z = zero,

(4) DS = decreasing slow, and

(5) DF = decreasing fast.

401. Hence, the helmsman's course-keeping behaviour is described by:

1. If ε is PVS	and $\dot\psi$ is any	then δ is PVS else
2. If ε is PVVS	and $\dot\psi$ is FI or IS or Z	then δ is PVVS else
3. If ε is PVVS	and $\dot\psi$ is DS	then δ is Z else
4. If ε is PVVS	and $\dot\psi$ is DF	then δ is NVVS else
5. If ε is Z	and $\dot\psi$ is FI	then δ is PVVS else
6. If ε is Z	and $\dot\psi$ is IS or Z or DS	then δ is Z else
7. If ε is Z	and $\dot\psi$ is DF	then δ is NVVS else

8. If ε is NVVS	and $\dot{\psi}$ is DF	then δ is PVVS else
9. If ε is NVVS	and $\dot{\psi}$ is DS	then δ is Z else
10. If ε is NVVS	and $\dot{\psi}$ is FI or IS or Z	then δ is NVVS else
11. If ε is NVS	and $\dot{\psi}$ is any	then δ is NVS

402.The output fuzzy set is then defined by:

$$\mu(\delta_c) = \max_{n=1\ldots 11}\left[\min\left\{\mu_{\varepsilon_n}(\varepsilon), \max\left(\mu_{\dot{\psi}_{n_1}}(\dot{\psi}), \mu_{\dot{\psi}_{n_2}}(\dot{\psi}), \mu_{\dot{\psi}_{n_3}}(\dot{\psi})\right), \mu_{\delta_n}(\delta)\right\}\right] \quad (2)$$

403. Again, the centre of area method of producing a deterministic value is used to obtain the helmsman's control output.

History and Source

404. From a survey of relevant literature, it was found that there were a few validated helmsman models available with the exception of Veldhuyzen and Stassen (1977) and Mort and Leonard (1982). Indeed, a review by Sutton and Towill (1986) revealed a general lack of man-machine systems theory being applied to marine problems. It was also discovered that since the proposition of fuzzy set theory by Zadeh (1965) numerous investigations have been carried out using the concepts in a variety of scientific disciplines. Surprisingly, however, little work has been performed using the theory to describe and analyse human operators involved in manual control tasks.

405. Hence, the work undertaken by Sutton and Towill (1987, 1988) at the Royal Naval Engineering College, Plymouth, UK and the University of Wales Institute of Science and Technology, Cardiff, UK, helps simultaneously to alleviate the lack of fuzzy set theory in man-machine applications and the dearth of man-machine systems theory being applied to marine problems.

406. For the reader interested in an introduction to fuzzy set theory, reference should be made to Zadeh (1973), Kochen (1975), Tong (1977), Zadeh (1984), Sutton and Towill (1985), and Zadeh (1988).

Product and Purpose

407. Both models, when performing in their respective modes, produce time-histories of their control action and statistical data regarding the performance achieved.

408. By using these models in closed loop simulation studies, a ship designer can make a better assessment of the handling qualities of a proposed vessel.

When used

409. At the systems design stage.

Procedures for Use

410. Each model is in a self-contained simulation package. Thus, the operator is only required to input the desired course and level of sea state disturbances.

Advantages

411. The rule-based structure of the models gives a more comprehensive understanding of the control behaviour of a helmsman.

Limitations

412. Relatively large digital computer facilities are required to run the simulation packages.

Application Examples

413. See references.

Technical Details

414. The models were simulated using a Control Data Cyber 180/840 mainframe digital computer using a FORTRAN 77 compiler running under a VE system.

References

Kochen, M. (1975). Applications of fuzzy sets in psychology. In L. A. Zadeh, K. S. Fu, K. Tanaka, & M. Shimura (Eds.). Fuzzy sets and their application to cognitive and decision processes. London: Academic Press.

Mort, N. & Leonard, C.H. (1982). Modelling the helmsman in a ship system. Proc. 4th International Symposium on Ship Automation. Genova.

Sutton, R. and Towill, D.R. (1985). An introduction to the use of fuzzy sets in the implementation of control algorithms. Journal of the Institution of Electronic and Radio Engineers, 55, (10) 357 - 367.

Sutton, R. and Towill, D.R. (1986). The helmsman as a man-machine element. Journal of Navigation, 39, (1) 49 - 65.

Sutton, R. & Towill, D.R. (1987). A fuzzy model of the helmsman performing a course-keeping task. Applied Ergonomics, 18, (2) 137 - 145.

Sutton, R. & Towill, D.R. (1988). Modelling the course-changing control behaviour of a helmsman using fuzzy sets. In J. Patrick & K.D. Duncan (Eds.). Training, human decision making, and control. Elsevier Amsterdam: North-Holland.

Tong, R.M. (1977). A control engineering review of fuzzy systems. Automatica, 13, 559 - 569.

Veldhuyzen, W. & Stassen, H.G. (1977). The internal model concept: An application to modelling human control of large ships. Human Factors, 19, (4) 367 - 380.

Zadeh, L.A. (1965). Fuzzy sets. Information and control, 8, 338 - 353.

Zadeh, L.A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. on Syst., Man, and Cybern, 3, (1) 28 - 44.

Zadeh, L.A. (1984). Making computers think like people. IEEE Spectrum, 21, (8) 26 - 32.

Zadeh, L.A. (1988). Fuzzy logic. IEEE Computer, 21, (4) 83 - 93.

3.3.16 An Integrated Pilot Model

Summary Description

415. An integrated pilot model was proposed and validated by Zacharias (1990) that includes the active perception and control of visually-guided self-motion. The optimal control model (OCM) of the human pilot is used to model the operator's information-processing and control activities. A visual cueing model (VCM) is added comprising four perceptual sub-models: (1) instrument cueing, (2) linear perspective cueing, (3) textural cueing, and (4) preview cueing.

History and Source

416. The integrated pilot model is founded in modern estimation and control theory, and is motivated by past and current studies of piloted flight control, particularly low-level terrain-following (TF) flight, an intense and demanding visually-driven task.

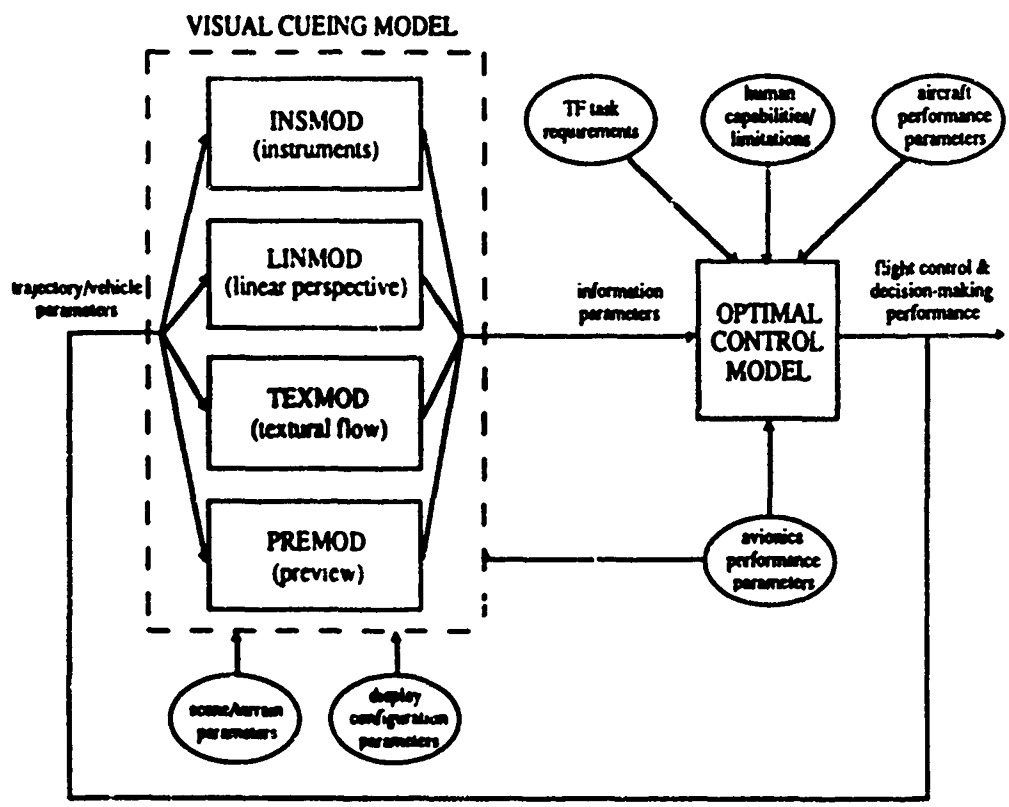


Fig. 3.21 Architecture of the integrated pilot model.

Product and Purpose

417. The integrated pilot model is an extension of the optimal control model (section 3.3.10) of the human pilot. The sub-models described above can serve as the basis for

developing an integrated estimation/control model for understanding visually-guided active self-motion. Figure 3.21 illustrates an overall architecture for such a model, specialized for the terrain-following environment. The optimal control model (OCM) is used to model the operator's information-processing and continuous control activities. A visual cueing model (VCM), comprised of the four perceptual sub-models just described, is used to model the operator's interaction with his display environment, and to model his resulting perceptual performance.

418. Both the OCM and the VCM are required to fully specify operator behavior in any realistic visually-driven egomotion task. Because the OCM works at the informational level, the VCM must provide the critical interface between the external-world display attributes, and the internal-world informational variables. The VCM, in effect, serves to transform the explicit display variables, which are defined by the display geometry and the physics of any intervening display technology, to the implicit informational variables, which are defined by the task at hand and the psychophysics of the human operator.

419. As has been indicated, the VCM appropriate to terrain-following flight is comprised of four sub-models, each of which accounts for different display types and configuration parameters, as described earlier. Thus, INSMOD models simple instrument cueing, and can account for such display factors as instrument resolution and dynamics. Likewise, LINMOD models complex linear perspective cueing, and can account for a variety of geometric relations between observer and scene (or computer-generated scene). TEXMOD models dynamic textural cueing, and can account for such factors as observer motion relative to solid objects in the visual world, whether real or display-generated. Finally, PREMOD models preview cueing and can account for the pilot's processing of future flight path information as seen on a terrain profile display, or as viewed out the window.

420. Figure 3.21 also shows how the OCM serves to integrate the VCM-generated informational variables with the other task-relevant factors, to support the prediction of the pilot's overall task performance. The use of the OCM allows one to account for the pilot's fundamental information-processing capabilities and limitations, and integrate these internal factors with critical external factors, such as the display characteristics, the flight task requirements, the aircraft's performance and response, and the capabilities of the supporting avionics. The overall integration of these factors within the structure provided for by the model then allows one to predict task-specific continuous flight control performance.

421. A submodel of the OCM deserving separate mention is an instrument cueing model (INSMOD) which models the operator's processing of visual cues presented via conventional instrument displays. In the flight environment, such displays include dedicated pointer/bar displays (e.g., airspeed) and programmable alphanumeric displays (e.g., alphanumeric data on a head-up display (HUD)), as well as simple tracking displays (e.g., a dedicated ILS error display, or a computer-driven HUD pipper). In short, any time a display output is functionally related to the variable driving it, in a simple fashion, it can be considered an "instrument". Hence INSMOD can be used to model the pilot's processing of the associated information. The model assumes that for simple instruments the pilot sees the displayed variable and its time rate-of-change.

422. The linear perspective visual cue model (LINMOD) can also be directly integrated into the overall model structure (Zacharias & Levison, 1980). The basic assumption underlying this model is that the perspective cues associated with a visual scene can be fully represented by an idealized line drawing of that scene, and furthermore, that the component line elements can be functionally related in a direct manner to the observer's positional and

attitudinal state with respect to the scene objects.

423. A more recent development is a visual textural cueing model (TEXMOD) for application to the dynamic analysis and modelling of scenes dominated by texture cues (Zacharias, Caglayan, & Sinacori, 1985 a, b). The model was developed and implemented to support the simulation and understanding of the pilot's processing of optic flow-field cues, during low-level terrain-following flight. The model is predicated on the notion that the pilot makes noisy, sampled measurements on the spatially-distributed optic flow-field surrounding him, and, on the basis of these measurements, generates estimates of his own linear and angular velocities with respect to the terrain surface. A subsidiary but significant output of the model is an "impact time" map, an observer-centered spatially-sampled scaled replica of the viewed surface.

424. The preview cueing model (PREMOD) can account for the pilot's processing of future flight path information. In most visually-driven active locomotion tasks, the operator is typically provided with information regarding future changes in the upcoming path. For example, in the terrain-following task, the pilot is provided information regarding the future desired flight path via such features as terrain surface curvature and roadway edges. The basic approach to modeling this processing of previewed path information rests on a transformation from the continuous future-time curve/surface domain to the discrete current-time parametric domain, via a model fit to the previewed path. That is, by fitting a parametric curve/surface model to the previewed path, the future path information is transformed to current-time estimates of the parametric model. These current-time estimates then serve as the basis for subsequent processing by the pilot model to support current-time discrete decisions and continuous control actions. This type of preview cueing is modeled by assuming that the pilot sees a curve which he internally models as an Nth order polynomial. The pilot is considered to take M noisy measurements of the previewed curve with which he generates a weighted least squares estimate of its parameters.

When used

425. This model may be used in the design of experiments or in the development of terrain-following displays.

Procedures for Use

426. Procedures are similar to those for the OCM.

Advantages

427. The visual cueing model permits the treatment of a broader range of cues than previous OCM implementations.

Limitations

428. Not available.

Application Examples

429. In this section, three model applications are summarized to demonstrate the model's utility in accounting for a range of visually-driven perception and control egomotion

tasks. In one study, Zacharias et al. (1985 a, b) compared the estimation performance of one component of the VCM, the TEXMOD sub-model, with data generated in a passive psychophysics experiment. Subjects were flown over flat terrain decorated by a uniformly random array of luminous dots. Using the TEXMOD submodel, noisy observation of the visual flow was simulated by additive measurement noise corrupting the line-of-sight (LOS) measurement vector couple. This was in remarkable agreement with a demonstrated 10 % figure obtained by independent threshold measurements, and strongly supports the direct functional linkage provided by the TEXMOD sub-model, linking a basic perceptual threshold with observed estimation performance on an egomotion-related task.

430. In a second study, Zacharias (1985) again compared model and data, this time looking at overall active flight task performance. The main task of the experiment was to maintain constant assigned altitude, while "flying" over flat terrain in the presence of a vertical wind gust. No lateral flight control was required of the subject, and the task was thus a single-axis disturbance regulation task. The model-based analysis demonstrated an ability to closely match performance score and frequency response data, across a wide range of experimental conditions. Most of the model matches were within a fraction of an across-subject standard deviation.

431. In a third study, a range of display enhancements for terrain-following flight were evaluated, via a combined analytic modeling and expefimental simulation effort; study results are in Brun and Zacharias (1986). An undulating random-appearing terrain drove an on-board guidance system, which, in turn, generated a vertical plane desired flight path to be followed by the pilot. The path was generated so as to maintain an approximately constant altitude above the terrain, while avoiding high-acceleration maneuvers. The simulation also had provisions for lateral path control and display, although they were not used in the experimental program. A range of terrain-following displays were evaluated to assess their impact on pilot performance and flight control strategy The simulation data was well-matched by the model points.

Technical Details

432. Not available.

References

Baron, S. (1987): Application of the optimal control model to assessment of simulator effectiveness. In W. B. Rouse (Ed.). Advances in man-machine systems research. Greenwich, CT: JAI Press.

Brun, H. M., & Zacharias, G. L. (1986). Model-based methodology for terrain-following display design. (Report No R8603). Cambridge, MA: Charles River Analytics.

Zacharias, G. L. & Levison, W. H. (1980). Development of a model for the use of extra-cockpit visual cues. (BBN Report No. 4562). Cambridge, MA: BBN Laboratories.

Zacharias, G. L. (1985). Modelling the pilot's use of flight simulator visual cues in a terrain-following task (Report No. R8505) Cambridge, MA: Charles River Analytics.

Zacharias, G. L., Caglayan, A. K., & Sinacori, J. B. (1985 a). A model for visual flo-field cueing and self-motion estimation. IEEE Transactions on Systems, Man and Cybernetics, 15, (3) 385-389.

Zacharias, G. L., Caglayan, A. K., & Sinacori, J. B. (1985 b). A visual cueing model for terrain-following-applicatons. AIAA Journal of Guidance, 8, (2) 201-207.

Zacharias, G. L. (1989): Motion-based state estimation and shape modeling. In Elkind, J. I., Card, S. K., Hochberg, J., and Huey, B. M. (Eds.). Human performance models for computer-aided engineering. Washington, DC: National Academy Press.

Zacharias, G. L. (1990): An estimation / control model of egomotion. In R. Warren & A. H. Wertheim (Eds.). Perception & control of self-motion. Hillsdale, NJ: L. Erlbaum Associates.

Future Needs

433. Not available.

CHAPTER 4

MULTI-TASK MODELS

4.1 INTRODUCTION: MULTI-TASK MODELS

434. The growing interest in multiple task performance and related models depends on many factors. Due to increasing automation and the use of advanced information technology, the human's function in vehicle and process control is shifting from a direct and continuously active involvement towards a supervisory control structure, where man-machine interaction is exercised through the mediation of a computer. Supervisory control indicates a hierarchy or coordinated set of human activities that includes initiating, monitoring, detecting events, recognising, diagnosing, adjusting, and optimising processes in systems that are otherwise automatically controlled. In many of these situations, the human operator has to accomplish several tasks at the same time with his attention and effort appropriately allocated among the tasks. Industrial process monitoring, multiprocess scheduling, aircraft piloting and air traffic control are among the more obvious examples. It appears that the limits of human decision making capability become obvious as the demand or complexity of the tasks increases. Most of the literature that relates to this section fits into the category of multiple task decision making. Thus the broad spectrum of decision making seems to be a primary characteristic of human operator activities, especially in the field of vehicle and process control.

4.2 OVERVIEW AND RECOMMENDED REFERENCES: MULTI-TASK MODELS

435. There are two ways to develop integrative models (i.e., models representing more than a single task-component) of the human operator controlling large-scale systems: bottom-up and top-down. The bottom-up approach starts with component models, findings, and data (e.g., eye movements, perceptual and central processes, performing manual actions, etc.) and attempts to combine them into an overall or aggregate model. Here the human operator simulator (HOS), section 4.3.7, is an appropriate example that proposes to systematise the process of model construction. The bottom-up approach has been successful primarily for modelling limited task segments. The difficulties inherent in integrating component research and in aggregating task-elements were pointed out by various authors (Pew et al., 1977; Pew and Baron, 1983; Sheridan, 1987). Perhaps the most difficult challenge of the bottom-up approach is the formulation of a suitable task taxonomy for deriving the components of human performance. The top-down approach starts with an overall system concept and progressively adds structure and defines variables implying appropriate data requirements. This approach reduces the problem of defining a suitable task taxonomy, but in its place a critical need arises for well-defined dimensions and objective functions which the man-machine system is to minimise. The PROCRU-model (section 4.3.6) involves features of a top-down approach, such as the normative model structure confirmed by the optimal control model (OCM).

436. The model summaries which follow borrow heavily from the original sources given in the lists of references. In many cases direct quotations are not indicated.

References

Baron, S. & Levison, W.H. (1980). The optimal control model: Status and future issues. Proceedings, IEEE Conf. on Syst., Man, and Cybern. New York: IEEE.

Baron, S., Kruses, D. S. & Huey, B. M. (Eds.) (1990). Quantitative modelling of human performance in complex, dynamic systems. Washington, DC: National Academy Press.

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In W.B. Rouse (Ed.). Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Bell, D. (Ed.) (1988). Decision making: Descriptive, normative, and prescriptive interactions. Cambridge, MA: Harvard Business School.

Card, S.K., Moran, T.P. & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: L. Erlbaum.

Chubb, G.P., Laughery, K.R. & Pritsker, A.A.B. (1987). Simulating manned systems. In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Dutton, J.M. & Starbuck, W.H. (1971). Computer simulation of human behavior. New York: J. Wiley.

Edwards, W. (1987). Decision making. In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Elkind, J. I., Card, S. K., Hochberg, J., & Huey, B. M. (Eds.) (1989). Human performance models for computer-aided engineering. Washington, DC: National Academy Press.

Hess, R.A. (1987). Feedback control models. In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Levis, A.H. (1984). Information processing and decision making organizations: A mathematical description. Large-scale systems, 7, 151-163.

McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

McRuer, D.T. (1980). Human dynamics in man-machine systems. Automatica, 16, 237-253.

Meister, D. (1985). Behavioral analysis and measurement methods. New York: J. Wiley.

Miller, R.A. (1984). A systems approach to modelling discrete control performance. In W.B. Rouse (Ed.). Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Moray, N. (1986). Monitoring behavior and supervisory control. In K.R. Boff, L. Kaufmann & J.P. Thomas (Eds.). Handbook of perception and human performance (Vol. 2). New York: Wiley.

Patrick, J. & Duncan, K.D. (Eds.)(1988). Training, human decision making and control. Amsterdam: North-Holland.

Pew, R.W., Baron, S., Feehrer, C.E. & Miller, D.C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation. (BBN Report No. 3446). Cambridge, MA: BBN Laboratories.

Pew, R.W. & Baron, S. (1983). Perspectives on human performance modelling. Automatica, 19 (6), 663-676.

Rouse, W.B. (1980). Systems engineering models of human-machine interaction. New York: North Holland.

Rouse, W.B. (1981). Human-computer interaction in the control of dynamic systems. Computing surveys, 13 (1), 71-99.

Sheridan, T.B. & Ferrell, W.R. (1974). Man-machine systems: Information, control and decision models of human performance. Cambridge, MA: MIT Press.

Sheridan, T. B. & Johannsen, G. (Eds.) (1976). Monitoring behavior and supervisory control. New York: Plenum Press.

Sheridan, T.B. (1984). Supervisory control of remote manipulators, vehicles and dynamic processes: Experiments in command and display aiding. In W.B. Rouse (Ed.). Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Sheridan, T.B. (1984). Research and modelling of supervisory control behavior. Washington, DC: National Academy Press.

Sheridan, T.B. (1987). Supervisory control. In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Stein, W. (1989). Models of human monitoring and decision making in vehicle and process control. In McMillan, G R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

von Winterfeldt, D. & Edwards, W. (1988). Decision analysis and behavioral research. New York: Cambridge University Press.

Warren, R. & Wertheim, A. H. (Eds.) (1990): Perception & control of self-motion. Hillsdale, N J: L. Erlbaum Associates.

Wickens, C.D. (1986). The effects of control dynamics on performance. In K.R. Boff, L. Kaufmann, J.P. Thomas (Eds.). Handbook of perception and human performance (Vol. 2). New York: J. Wiley.

Wohl, J.G., Entin, E.E., Kleinman, D.L. & Pattipatti, K. (1984). Human decision processes in military command and control. In W.B. Rouse (Ed.). Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Woods, D.D., O'Brien, J.F., Hanes, L.F. (1987). Human factors challenges in process control: The case of nuclear power plants. In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

4.3 MODEL SUMMARIES: MULTI-TASK MODELS

4.3.1 A Model of Human Decision Making in Multiple Process Monitoring

Summary Description

437. A model of human display monitoring has been developed and validated in a laboratory situation by Greenstein and Rouse (1982). In that task situation the human operator simultaneously monitors multiple dynamic processes for action evoking events. The processes may differ in priority and the human may be unable to attend to all processes simultaneously, instead having to allocate his attention among processes.

History and Source

438. Not available.

Product and Purpose

439. Greenstein and Rouse (1982) have employed the display monitoring model for describing human event detection in multiple process monitoring situations. The task which they considered is illustrated in Figures 4.1 and 4.2. Subjects simultaneously viewed the sampled outputs of nine second-order dynamic processes. Their task was to detect changes in the signal-to-noise ratio characterised by the process outputs becoming increasingly noisy. Their instructions were to detect changes as quickly as possible while also avoiding false alarms. Considering Figure 4.1, events (i.e., the onset of the decreasing signal-to-noise ratio) had occurred in processes 1, 2, 3, and 8 at times 130, 113, 92, and 106, respectively. At the point at which this display was generated, the subject had only detected the event in process 3 while also incorrectly responding to process 5. Figure 4.2 illustrates the display 10 time units later (one update). The dashed vertical lines indicated to the subject the last point of response to each process. In studying this event detection task, Greenstein and Rouse were primarily interested in developing a model for situations where knowledge of the dynamics of the processes was unavailable and thus one had to observe the human to determine how the task could be performed.

When Used

440. The model seems to have considerable practical implications, especially in the realm of computer-aided decision making. Consider a situation in which a human must simultaneously monitor many processes. Further, assume that a lack of knowledge of the dynamics of the processes (e.g., a chemical plant) as well as the presence of time-varying priorities and costs preclude direct automation of the monitoring tasks. Using the event detection model described here, a computer system could be designed to "watch" the human operator and hence be capable of providing back-up decision making if the human were to become overloaded.

Procedures for Use

441. Not available.

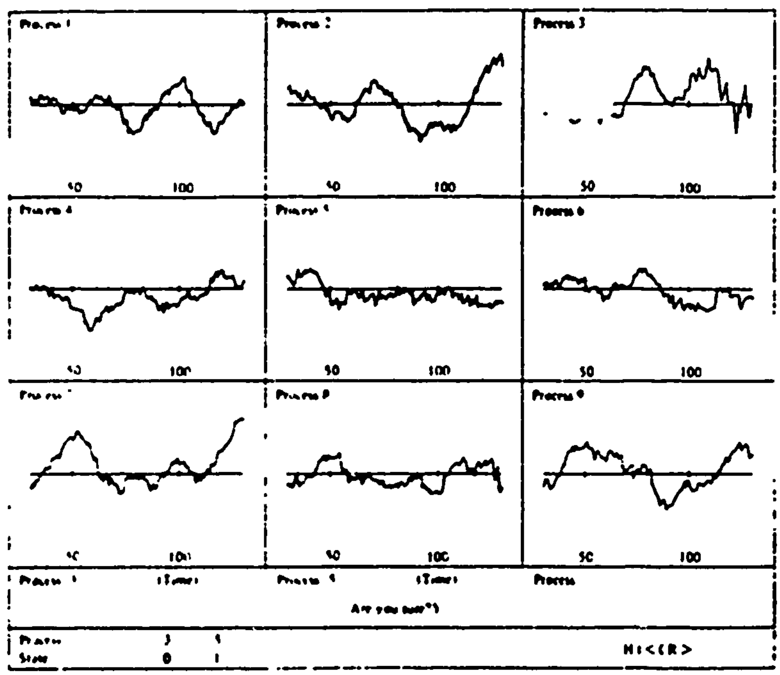


Fig. 4.1 The multiple process monitoring situation.

Advantages

442. The display monitoring model has been shown to provide a good fit to data obtained using a specific multiple process monitoring situation. The use of discriminant analysis to generate event probabilities permits the model to be used in situations in which explicit models of the processes being monitored are unavailable or in which relevant features are not related in a straightforward manner to the dynamics of the process. It also allows the model to be applied with relative ease to situations in which specific event probability estimation algorithms are available.

Limitations

443. The model has only been tested in experimental situations.

Application Examples

444. To test the usefulness of the pattern recognition model, an experiment was conducted. As they performed the task, subjects kept written "logs" of their actions and descriptions of what they were doing. Based on these logs, four features were selected for use in the model. The first feature involved the magnitude changes in the sequence of recent process outputs. The second feature entailed the presence of reversals in direction in this sequence (i.e., changes of slope from positive to negative or vice versa. The third feature was based on the simultaneous occurrence of large-magnitude changes and reversals. Finally, the fourth feature was a very local measurement of magnitude changes in the most recent four

points in the sequence of process outputs. In calculating average changes in magnitude, average number of reversals, and so on, an exponential averaging method was used to accommodate the fact that events became more pronounced as they evolved.

445. The model was tested by allowing it (i.e., the computer) to "watch" each subject during one of his or her trials. From these observations data were taken and processed. The model's performance was then compared to that of the subject on another trial (i.e., one on which the model had not been trained). It was found that the model did quite well in terms of matching a subject's average time to event detection and number of correct detections, although it was somewhat more conservative than subjects in terms of false alarms. In applying the model to this situation, very little fitting of parameters to data was required. Some parameters of the model were fixed at values considered to be intuitively reasonable. Features were selected in similar fashion, although the comments of experimental subjects performing decision making tasks within the situation were also instructive.

Technical Details

446. Not available.

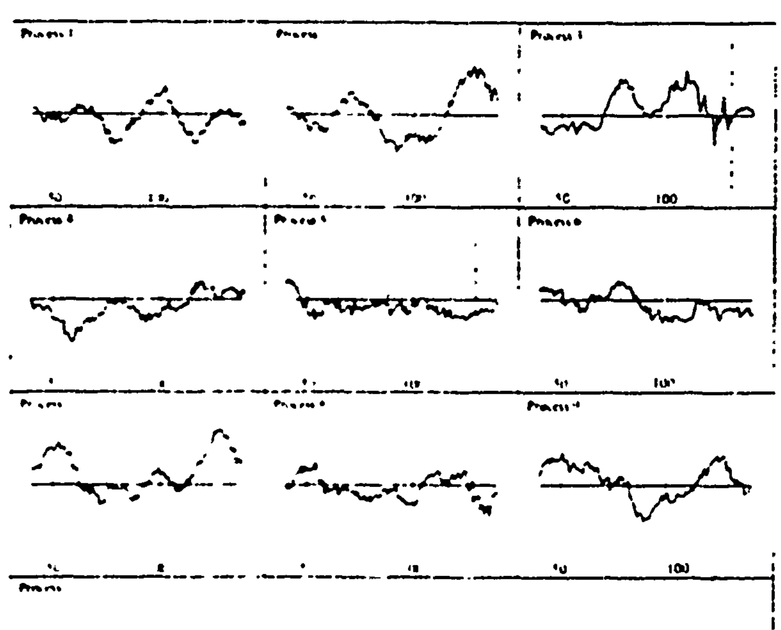


Fig. 4.2 An updated display

References

Greenstein, J. S., W.B. Rouse (1982). A model of human decision making in multiple process monitoring situations. IEEE Trans. Syst., Man, and Cybern., 12, (2) 182-193.

Rouse, W.B. (1980). Systems engineering models of human-machine interaction. New York: North-Holland.

Rouse, W.B. (1981). Human-computer interaction in the control of dynamic systems. Computing Surveys, 13, (1) 71-99.

Future Needs

447. Not available.

4.3.2 A Model of the Human Controller in Combined Continuous and Discrete Tasks

Summary Description

448. A model of the human supervisor in a multi-task situation involving continuous control and discrete tasks has been proposed by Govindaraj and Rouse (1981). The model has a number of parameters that can be varied for matching the experimental results. The parameters are (1) ratio of weights on control to weights on error, (2) ratio of nominal weights on control to weights on control over discrete task intervals, and (3) threshold on changes in control. After some preliminary trials, the threshold for changes in control could be fixed which resulted in performance similar to that of the subjects.

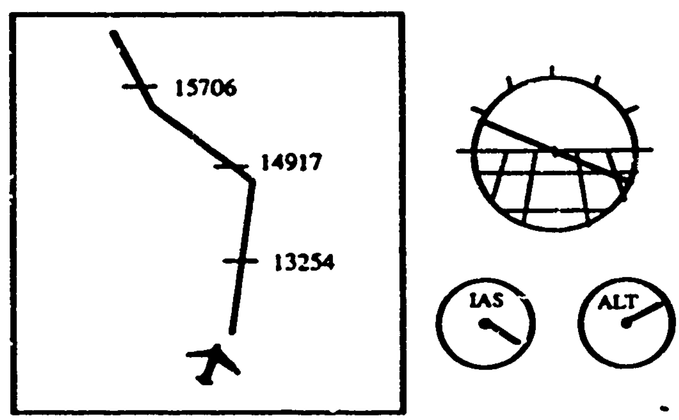


Fig. 4.3 Multi-task flight management situation.

History and Source

449. The model developers were trying to answer the question of how a pilot schedules discrete tasks (e.g., talking with air traffic control, taking radio fixes, and so on) while also performing a continuous control task. This task presents difficulties for the usual optimal control models because it is typical for the pilot to stop controlling (i.e., stop moving the control stick) while performing some types of discrete tasks. As most manual control models produce continuous outputs, a new formulation was needed.

Product and Purpose

450. This model is concerned with the human operator when he must perform different types of tasks simultaneously. An appropriate example is that of flying a modern airliner. The tasks expected of a pilot can be broadly subdivided into continuous control tasks, where he uses various control devices to alter the flight conditions, and discrete tasks such as checking subsystems during various phases of flight, communicating with air traffic controllers, and taking radio fixes.

451. Optimal control theory was used to develop an analytical model for the problem of controlling a system so as to move along a reference trajectory. Preview of the reference trajectory was assumed to be available for some distance (or equivalently, time). This is the

planning horizon in which the discrete tasks must be inserted at appropriate intervals. A well-trained, well-motivated human operator is assumed to keep the errors from the reference trajectory to a minimum. He is also expected to impose a penalty on the control action.

452. The important feature in scheduling control and discrete tasks is the possibility of directly influencing the control through the corresponding weights. When it is desired to insert a discrete task somewhere in the planning horizon, the control weighting is increased in that interval, resulting in a reduced control value. If this is done for the entire planning horizon (i.e., altering the weights in intervals of desired discrete task activity), this has the desirable effect of redistributing control activity to keep the overall cost at an optimal value.

When Used

453. The model seems to be useful in evaluating displays where the map, or in general, the future reference is known for a certain distance. If discrete task characteristics are known, the amount of time required for the discrete tasks and when they should be performed could be determined. From the perspective of computer aiding, where responsibilities are shared between the computer and the human, the model could be valuable for predicting how different allocations of tasks affect control task performance.

Procedures for Use

454. Not available.

Advantages

455. When combined with the queuing model developed by Rouse (1977), Walden and Rouse (1978), and Chu and Rouse (1979), the model presented here would allow for predictions of overall performance for both continuous control and discrete tasks.

Limitations

456. The model has only been tested in experimental situations.

Application Examples

457. An experiment was designed to test the validity/appropriateness of the model in the multi-task flight management situation shown in Fig. 4.3. The display was a standard graphics terminal. For the aeroplane dynamics used, the updates appeared to move smoothly without flicker, and without long persistence.

458. The control task involved flying over a map at a constant altitude and airspeed. The subject controlled the aileron angle using a joystick. A "track up" display was used where the map to be flown moved past a fixed aeroplane symbol. The reference trajectory for 53 s into the future was available for preview. An artificial horizon with a moving wing over a fixed horizon was used to indicate the attitude. The airspeed indicator showing the indicated air speed (IAS) and the altimeter (ALT) were included only to provide a measure of completeness.

459. A number of discrete tasks appeared in the preview period. Associated with each discrete task was a starting position marked by a tic mark on the map. This also served as the deadline for completion of the preceding discrete task. After completion of service to a discrete task, the next one could not be started until the aeroplane symbol passed the starting position.

460. Discrete tasks were presented as data entry tasks. At various times five digit numbers appeared and moved along the map. Successful performance of these discrete tasks required correct entry of the numbers before they disappeared. A number would stay on the map until it was correctly entered, or the position on the map corresponding to that number disappeared from the display. Some or all of these digits could be entered at one time.

Technical Details

461. Not available.

References

Chu, Y.-Y. & Rouse, W. B. (1979): Adaptive allocation of decision making responsibility between human and computer in multi-task situations. IEEE Trans. on Systems, Man, and Cybernetics, 9, (12) 769-778.

Govindaraj, T. & Rouse, W. B. (1981): Modelling the human controller in environments that include continuous and discrete tasks. IEEE Trans. on Systems, Man, and Cybernetics, 11, (6) 410-417.

Greenstein, J.S. & Rouse, W. B. (1982). A model of human decision making in multiple process monitoring situations. IEEE Trans. Syst., Man, and Cybern., 12, (2) 182-193.

Rouse, W. B. (1977): Human-computer interaction in multi-task situations. IEEE Transactions on Systems, Man, and Cybernetics, 7, (5) 384-392.

Rouse, W.B. (1980). Systems engineering models of human-machine interaction. New York: North-Holland.

Rouse, W.B. (1981). Human-computer interaction in the control of dynamic systems. Computing Surveys, 13, (1) 71-99

Walden, R. S. & Rouse, W B. (1978): A queuing model of pilot decision making in a multi-task flight management situation. IEEE Trans. on Systems, Man, and Cybernetics, 8, (12) 867-875.

Future Needs

462. It is desirable to test the model in more realistic situations. Multi-axis control tasks need to be considered where appropriate. It is possible that the methods developed here might be applicable without the need for any basic changes. Also, discrete tasks characterised by different priorities and different arrival and service times statistics would increase realism. Experimental situations similar to those reported in Walden and Rouse (1978) and Chu and Rouse (1979) might be used. Using training simulators is another possibility, where a variety of conditions could be tried.

4.3.3 A Model of Human Problem Solving in Dynamic Environments

Summary Description

463. Human performance in monitoring and controlling a large-scale system, such as a communication network, is considered. A model of performance in monitoring and controlling a simulated large-scale system (Figure 4.4) has been proposed by Zinser and Henneman (1988) and Henneman (1988) that explicitly incorporates three types of knowledge: system, contextual, and task knowledge. The first two types are represented by a network of frames; the third type is represented by a set of rules. The model compared very favourably to human performance in an experimental validation.

History and Source

464. The model presented in this summary is an extension of a conceptual model of human problem solving proposed by Rouse (1983). Rouse has suggested that problem solving takes place on three levels: (1) recognition and classification; (2) planning; and (3) execution and monitoring. Thus, when a problem situation develops, the first task is to detect that the problem exists and to categorise it (recognition and classification). An approach or plan to solving the problem must then be developed (planning), and finally, the plan must be implemented (execution and monitoring). The model is further characterised by its ability to make either a state- or a structure-oriented response, depending on both the system state and the human's level of expertise. The model assumes that humans have a preference for pattern-recognition solutions to problems - that is, humans prefer to make context-specific state-oriented responses to situations. Different extensions of the conceptual model of problem solving have been proposed by Hunt and Rouse (1984), Knaeuper and Rouse (1985), and Morris and Rouse (1985).

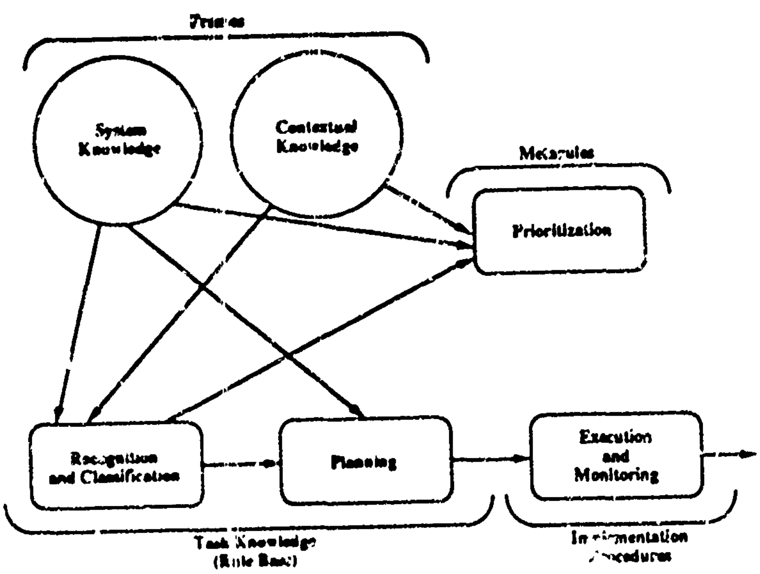


Fig. 4.4 Components of the conceptual model of human problem solving.

Product and Purpose

465. With this modelling approach, human performance in monitoring and controlling large-scale systems is considered. A simulated large-scale system, CAIN, was developed to provide an experimental vehicle of sufficient complexity to facilitate these human performance studies. CAIN is structured as a large hierarchical network that contains thousands of nodes. Customers travel through the system from a randomly selected source node to a random destination. Subjects monitor this system activity via a CRT display. When they detect a problem in the system (possibly due to a failure), subjects issue an appropriate command through a keyboard to correct and compensate for the abnormal situation. Ten different commands are possible. The overall objectives of the operator are to maximise the number of customers served and to minimise the time it takes for customers to travel between source and destination nodes.

466. A typical display of CAIN is shown in Figure 4.5. Because there are so many nodes in the network, it is not possible to display information about all nodes at one time. Thus nodes are grouped into relatively small networks called clusters. Human operators are restricted to viewing only one cluster at a time on the CAIN display. Clusters are grouped into hierarchy levels. The right side of the display shows information about one cluster of nodes, the middle left portion of the display presents a variety of subject-requested information about the system state, and the subject inputs control actions and information requests at the bottom left of the display. Each node in CAIN is referred to by a geographic label. Subjects can form associations or links between system parts, therefore, due to the existence of this contextual information.

467. In the context of this model, dynamic environment means that the human has to cope with the dynamics inherent in the physical world. The system to be controlled changes dynamically over time. Results of actions do not take place immediately and predictions of future system states are rather complicated. Knaeuper and Rouse's (1985) review of a variety of approaches to categorising the tasks an operator has to perform while controlling a dynamic process led to the choice of a classification scheme involving four general tasks, two or more of which may have to be performed simultaneously: (1) transition tasks, such as start-up, shut-down, take-off, and landing; (2) steady-state tuning; (3) detection and diagnosis of failures; and 4) compensation for failures.

468. To perform these tasks, the operator has to know (1) how the process will evolve if left alone; (2) what the effect will be of implementing control actions; and (3) what task is currently appropriate. Transition tasks are fairly proceduralized, although there need not be formal written procedures. A certain sequence of actions is often known, which will lead to the desired outcome. Steady-state tuning involves actions oriented toward optimising performance. In contrast to transition tasks, tuning tasks are "left-hand" or pattern-driven since the operator generally does not work towards a certain goal-state of the system. Failure detection and diagnosis will necessarily be performed in parallel to transition and tuning tasks. More specifically, failure detection is active at all times and does not interrupt transition or tuning tasks. Once an abnormal condition has been detected, then diagnosis may begin. It is possible that the diagnostic task could be either pattern or goal-driven. Failure compensation tends, in most cases, to be fairly proceduralized. Once the cause of a disturbance has been diagnosed then, if it is a familiar failure, an appropriate sequence of actions can be performed. However, if it is an unfamiliar failure, alternative approaches to failure compensation may have to be considered.

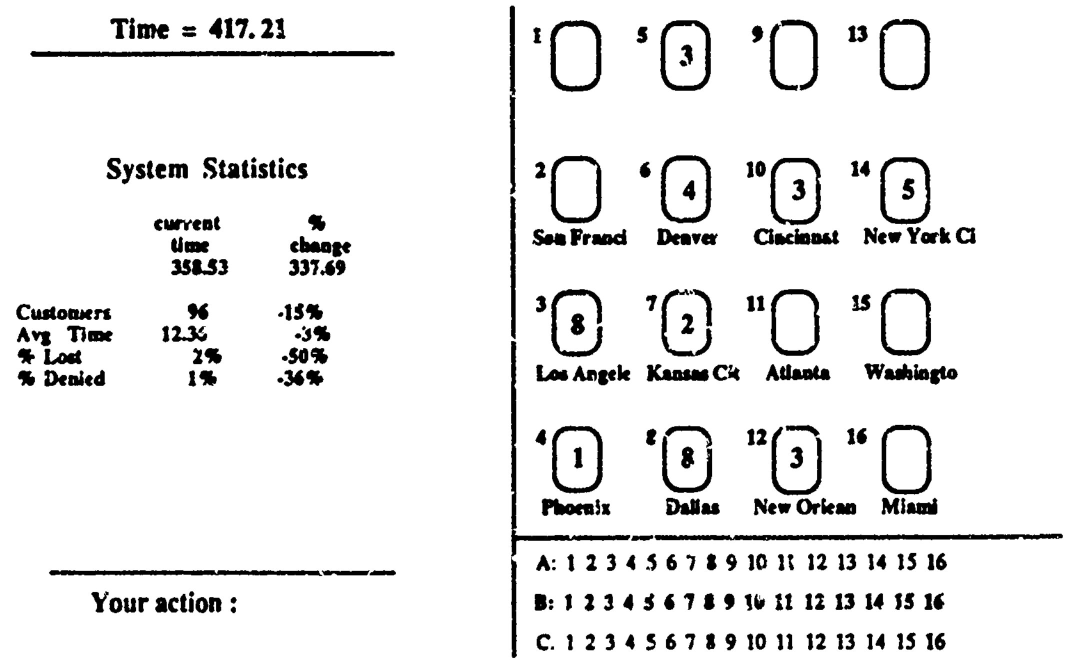


Fig. 4.5 Display of large-scale system CAIN.

469. The model proposed for the CAIN environment, MURRAY, is illustrated in Figure 4.4. MURRAY operates in the three stages of Recognition and Classification, Planning, and Execution and Monitoring. Situations are continually reevaluated as system states change due to the system dynamics or operator actions. An important feature of this task is that at any given time the human operator may have several different tasks that could be performed.

470. MURRAY's fidelity to human performance is dependent on the representation of three different types of knowledge needed to perform the task: system knowledge, contextual knowledge, and task knowledge. System knowledge and contextual knowledge are shown explicitly in Figure 4.4, while the task knowledge is embedded within the Recognition/Classification and Planning components. The Execution component of the model is realised by implementational procedures and the command that is issued.

471. The first type of knowledge, system knowledge, consists of information from CAIN about the current system state, for example, the number of customers waiting to be served in a city. Thus, the system knowledge of MURRAY is identical to the information presented on the CAIN display (Fig. 4.5). System knowledge is only accessed by the model's Recognition/Classification component and the Prioritisation mechanism. The system knowledge is structured as a hierarchical frame system. The frame of the highest structural level represents the cluster currently displayed by CAIN. Each city frame has several "slots" that contain such information as the number of customers waiting for service at the city and the average length of time they have been waiting.

472. The second type of knowledge, contextual knowledge, consists of information concerning the context of the system at a given time, such as locations of individual cities in the network and cities that have high loading and abnormal failure rates. Thus, contextual knowledge is augmented over time; as the model gains "expertise", the knowledge stored by this component will change. Contextual knowledge is represented by a network of context frames.

473. Finally, the third type of knowledge, task knowledge, represents the operator's behaviour in monitoring, problem solving, and failure detection. Task knowledge is represented as a production system. The operator's heuristics correspond to productions (or rules), while the operator's internal model of the system corresponds implicitly to metarules that organise the application of the explicit rules.

474. MURRAY contains 22 rules in its representation of task knowledge. These rules have a fixed syntax, and thus, they can be manipulated from outside the programme by a text editor. The set of rules is based on a combination of expert judgment and empirical evidence. Each rule consists of a situation and an action part made up of predicates.

475. An important part of MURRAY is the inference mechanism of the rule-based representation of the task knowledge. This mechanism determines the way that rules are applied and evaluated. The mechanism is implemented whenever the system state changes, i.e., whenever the model observes a set of new data from CAIN (as a reaction to a command issued by the operator or a dynamic change in the system). At this point, the condition predicates of all the rules are evaluated successively in the Classification component of the model.

When Used

476. Not available.

Procedures for Use

477. Not available

Advantages

478. To summarise, MURRAY proved to be a reasonable means of describing human behaviour in a complex monitoring and control task. Open-loop analysis of model performance indicated that the model consistently did as well as human operators. Closed-loop, action-by-action comparison of subject and MURRAY performance revealed a high degree of behavioural congruence. Thus, it appears that the structures and mechanisms present in the model produce quite similar behaviours to humans' structures and mechanisms used in performing this task. Nevertheless, it should be noted that the level of matching was not perfect. Both MURRAY and human operators appear to have different strengths that are useful in this environment: MURRAY is good at prioritizing tasks; the human operator is good at improvising flexible search strategies. Thus, a combination of the two could result in improved overall system performance. The next step in this research programme, therefore, was to implement a human performance aid based on MURRAY. Such an aid should provide cognitively plausible assistance to the human operator.

Limitations

479. Not available.

Application Examples

480. Experimental data were collected to validate MURRAY. Comparison of MURRAY and subject performance was done in two ways. First, an open-loop comparison was made in which subject performance was compared with MURRAY's performance. Second, a closed-loop analysis was performed. Subject data files were replayed concurrently with a version of MURRAY. Whenever a subject action was performed, MURRAY generated the action it would implement, along with a list of its other applicable rules. The subject's action was then implemented. This form of analysis allowed an action-by-action (or process) performance comparison to be made.

481. Three measures, mean customer sojourn time, number of customers served, and fraction of failures repaired, were used to assess subject performance. MURRAY out performed all of the subjects in terms of mean sojourn time and number of customers served. MURRAY ranked eighth, however, in terms of the third performance measure, fraction of failures repaired.

482. Nevertheless, by relying on its task description provided in the rule base, MURRAY resulted in uniformly excellent performance. Therefore, a model-based aid might be useful in providing the operator with procedural instructions; MURRAY could support the operator with additional or alternative strategies to monitor or control CAIN.

Technical Details

483. Not available.

References

Henneman, R. L. & Rouse, W. B. (1984 a). Measures of human performance in fault diagnosis tasks. IEEE Transactions on Systems, Man, and Cybernetics, 14 (1), 99 - 112.

Henneman, R L. & Rouse, W. B. (1984 b). Human performance in monitoring and controlling hierarchical large scale systems. IEEE Transactions on Systems, Man, and Cybernetics, 14 (2), 184 - 191.

Henneman, R. L. & Rouse, W. B. (1986). On measuring the complexity of monitoring and controlling large scale systems. IEEE Transactions on Systems, Man, and Cybernetics, 16 (2), 193 - 207.

Henneman, R. L. (1988). Human problem solving in dynamic environments. In W. B. Rouse (Ed.). Advances in man-machine systems research (Vol. 4). Greenwich, CT: JAI Press.

Hunt, R. M. & Rouse, W.B. (1984): A fuzzy rule-based model of human problem solving. IEEE Transactions on Systems, Man, and Cybernetics, 14 (2), 112 - 119.

Knaeuper, A. & Rouse, W. B. (1985): A rule-based model of human problem-solving behavior in dynamic environments. IEEE Transactions on Systems, Man, and Cybernetics, 15 (6), 708 - 719.

Morris, N. M. & Rouse, W. B. (1985). The effects of type of knowledge upon human problem solving in a process control task. IEEE Transactions on Systems, Man, and Cybernetics, 15 (6), 698 - 707.

Morris, N. M., Rouse, W. B. & Fath, J. L. (1985). PLANT: An experimental task for the study of human problem solving in process control. IEEE Transactions on Systems, Man, and Cybernetics, 15 (6), 792 - 798.

Rouse, W.B. (1983). Models of human problem solving: Detection, diagnosis, and compensation for system failures. Automatica, 19 (6), 613 - 625.

Rouse, W. B. & Hunt, R. H. (1984). Human problem solving in fault diagnosis tasks. In W. B. Rouse (Ed.). Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Zinser, K. & Henneman, R. L. (1988). Development and evaluation of a model of human performance in a large-scale system. IEEE Transactions on Systems, Man, and Cybernetics, 18 (3), 367 - 375.

Zinser, K. & Henneman, R. L. (1989). A model-based aid for monitoring and controlling a large-scale systems. IEEE Transactions on Systems, Man, and Cybernetics, 19 (2), 888 - 892.

Future Needs

484. Not available.

4.3.4 A Dynamic Decision-Making Model (DDM)

Summary Description

485. Pattipati, Kleinman and Ephrath (1983) have developed a dynamic decision making model (DDM) for predicting human task sequencing. It contains the same information processing structure as in the optimal control model (OCM). As in the original Siegel and Wolf models, situation assessment in the DDM involves estimation of the time available and time required for task completion. These variables are obtained from a memoryless transformation of the estimated system state variables. The DDM does differ from the other decision making models discussed here in one significant way, namely in the introduction of human randomness in the decision making algorithm itself. Though the practical value of including this randomness in performing system design and analysis may be argued, it is clear that such randomness in human decision making does exist. Moreover, in the context of the relatively simple paradigm to which the DDM was applied, the introduction of the stochastic choice axiom allows the DDM to be used to compute performance statistics analytically, rather than by Monte Carlo simulation.

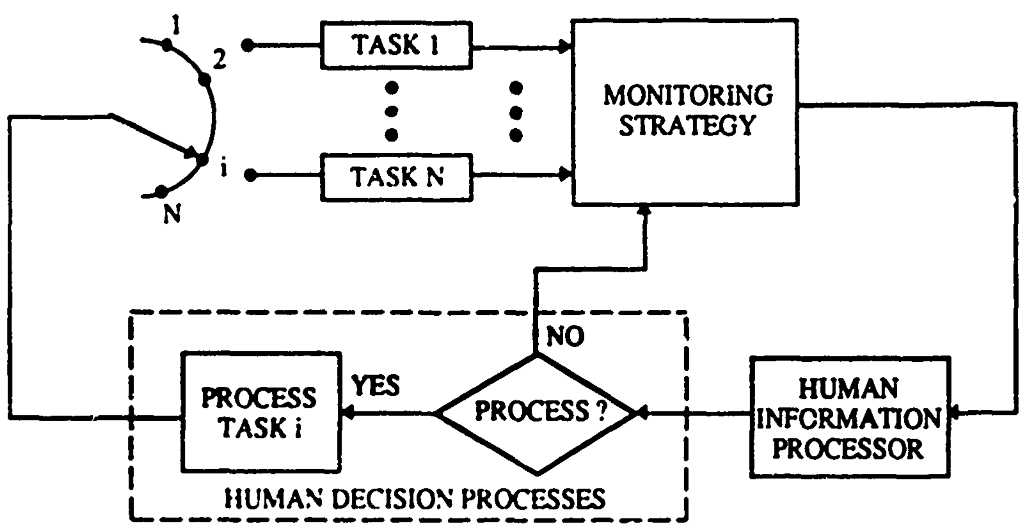


Fig. 4.6 Multi-task monitoring/decision making situation.

History and Source

486. This research was motivated largely by target selection problems. In this situation, targets of various types move across the display scopes of the human operator, vying for his attention. Since each target has different velocity and different distance to travel, the operator has variable length of time available to process a target before it disappears. Each target has different threat value and processing time requirements. The human is faced with the problem of sequencing targets dynamically.

Product and Purpose

487. The primary purpose of this approach is to gain a deeper understanding of human information processing and task selection procedures in dynamic multi-task environments. The approach has been to combine the results of a joint analytic and experimental programme into a normative dynamic decision model of human task sequencing performance. To this end a general multi-task paradigm was developed that retains the essential features of human task selection in a manageable yet manipulative context. Via this framework, they have studied the effects of various task related variables on the human decision processes. The model that has emerged from this effort forms a small but significant step toward human modelling in complex supervisory control systems.

488. The dynamic decision making model (DDM) uses the information processing structure of the OCM, extends the control theory approach to dynamic decision making and particularly addresses the problem of task selection in a dynamic multi-task environment. The experimental paradigm of Tulga and Sheridan (1980) was modified to provide an appropriate laboratory task for validing the DDM. In case of a situation with N independent tasks, the DDM includes a set of N independent estimator-predictor combinations.

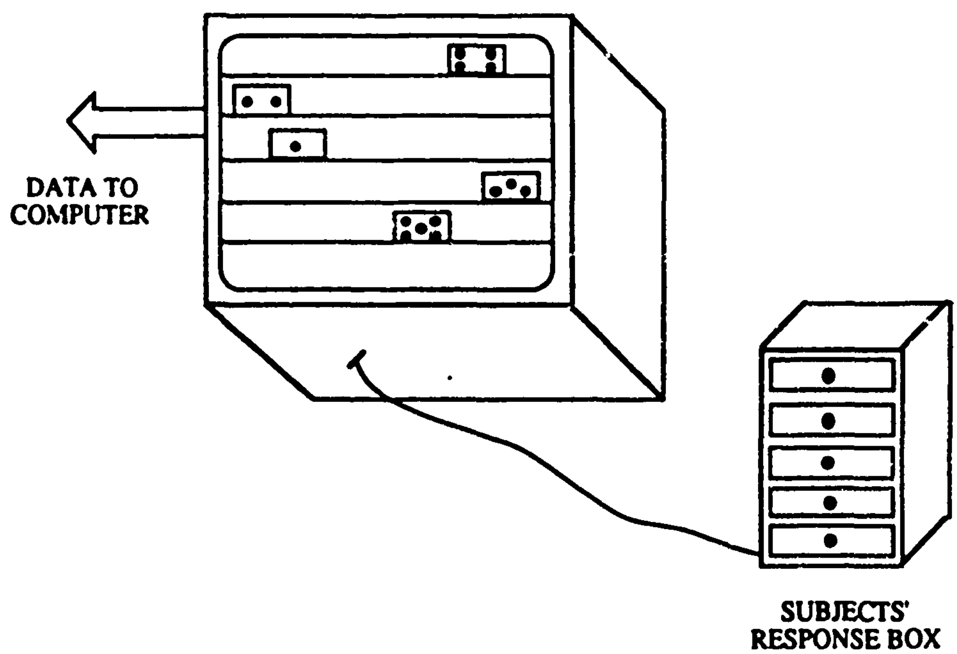


Fig. 4.7 Experimental apparatus.

489. Figure 4.6 shows the fundamental decision loop that is considered in this model. The human decision process involves 1) whether to process a task or gather more information (i.e., monitor), and 2) which of N tasks (N is time varying) to act upon, in order to maximise the system performance (e.g., maximise reward, minimise regret, etc.). The decision loop is dynamic in nature. As time evolves, tasks of different value, duration (processing time), and opportunity window demand the human's attention, while others depart. The opportunity windows shrink with time as the tasks approach their deadlines.

490. In the experiments conducted to develop the model, the subjects observed a computer screen on which multiple concomitant tasks were represented by moving rectangular bars. The bars appeared continually at the left edge of the screen and moved at different velocities to the right, disappearing upon reaching the right edge. Thus the screen width represented an opportunity window. In this experimental paradigm there could be, at most, a total of five tasks on the CRT screen, with a maximum of one on each line of any given time. The height (reward, value) of each bar was either one, two, or three units. The number of dots ($1<m<5$) displayed on a bar represented the time (in seconds) required to process the task. The subject could process a task in the "opportunity window" by holding down the appropriate push button as in Figure 4.7. By processing a task successfully, the subject was credited with the corresponding reward (r_i=1, 2, or 3), and the completed task was eliminated from the screen.

When Used

491. The DDM can predict various measures related to decision making performance, including task completion probability, and error probability. As in various other models, situation assessment in the DDM involves estimation of the time available and the time required for task completion.

Procedures for Use

492. Not available.

Advantages

493. A major contribution of the DDM work is the experimental validation of the model. By constraining the experimental paradigm to a situation that could be treated carefully in an experimental environment, it became possible to test model hypotheses with reasonable cost and control. These tests showed that the ideas underlying the DDM are essentially sound. They also provide further validation for the other control-theoretic models discussed herein.

Limitations

494. Not available.

Application Examples

495. The DDM has only been applied in experimental validations, but has proven quite accurate in that context.

Technical Details

496. The approach to modelling human decision behaviour parallells the optimal control model (OCM) of human response in spirit but not in form. In the OCM the control and information processing strategies are separable. Once an estimate of the system state is available, the linear feedback control law uses this estimate as if it were the true state. This type of separation has been found to be plausible in the present dynamic decision model. Analogous to the system state in OCM, the important concept of decision state is introduced in the DDM. The decision state variables are chosen to satisfy the axiomatic definition of a state,

i.e., it must provide the complete running summary of past actions (decisions). The joint density of the decision state is estimated from the information processor of the DDM, and provides sufficient information for the decision process.

References

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In: W.B. Rouse (Ed.), Advances in man-machine systems research (Vol. 1). Greenwich, CT: JAI Press.

Kleinman, D. L., Luh, P. B., Pattipati, K. R., & Serfati, D. (1990): Mathematical models of team distributed decision making. In: R. W. Swezey & E. Salas (Eds.), Teams: their training and performance. New York: Ablex.

Pattipati, K.R., Kleinman, D.L. & Ephrath, A.R. (1983). A dynamic decision model of human task selection performance. IEEE Trans. Syst., Man, and Cybern., 13, (3), 145-166.

Pew, R.W. & Baron, S. (1983). Perspectives on human performance modelling. Automatica, 19, (6), 663-676.

Rouse, W.B. (1981). Human-computer interaction in the control of dynamic systems. Computing Surveys, 13 (1), 71-99.

Tulga, M.K., Sheridan, T.B. (1980). Dynamic decisions and workload in multi-task supervisory control. IEEE Trans. Syst., Man, and Cybern., 10, (5) 217-232.

Future Needs

497. In the present experimental context the tasks are assumed to be independent, and the task values and velocities are constant as the bar moves across the CRT screen. However, in many realistic situations tasks are rarely independent and task attributes (e.g., value and velocity) may evolve in time, or they may vary as a function of human's decisions. Therefore future tests of DDM should consider more intricate task structure, such as those involving nonstationary task attributes, task dependency (e.g., precedence restrictions), and resource constraints.

498. A more realistic and challenging problem is the modelling of multiple DM's in distributed multi-task systems. Here, tasks arrive at each individual DM. An individual DM has to determine whether to keep an arriving task for himself or send it to someone else, and which task, if any, he should process. Thus the decision process requires the specification of a local routing strategy and a local sequencing strategy for each DM. The decision process is affected by the communication, information pattern at each DM, hierarchical structures, interhuman randomness, and variability, to name but a few.

4.3.5 A Model of Combined Monitoring, Decision-Making, and Control (DEMON)

Summary Description

499. DEMON is a multi-task model (including decision making, monitoring, and control tasks) for analysing the operational control of multiple remotely piloted vehicles (RPV). Control of each RPV represents a separate task, each with a payoff for maintaining errors within tolerance and for timely pop-ups and hand-offs. The operator's task is to monitor the trajectories and the estimated times of arrival (ETAs) of N vehicles, to decide if the deviation from desired flight path or the ETA error exceeds some tolerance threshold,and to correct respective paths by issuing appropriate control commands.

History and Source

500. The DEMON model was developed by Muralidharan and Baron (1979, 1980) and extended previous control-based approaches to a multi-task environment involving essentially discrete control decisions. The DEMON model is an important integrative step in the development of supervisory control models.

Product and Purpose

501. The essence of the DEMON approach is to characterise the operator limitations and the mission goals in a manner that allows one to predict operator strategies and overall system performance in human supervisory control of vehicles. For example, DEMON considers the detection of events not explicitly related to the system state variables and accounts for the time to complete discrete tasks. By means of a model extension, multi-operator situations as well

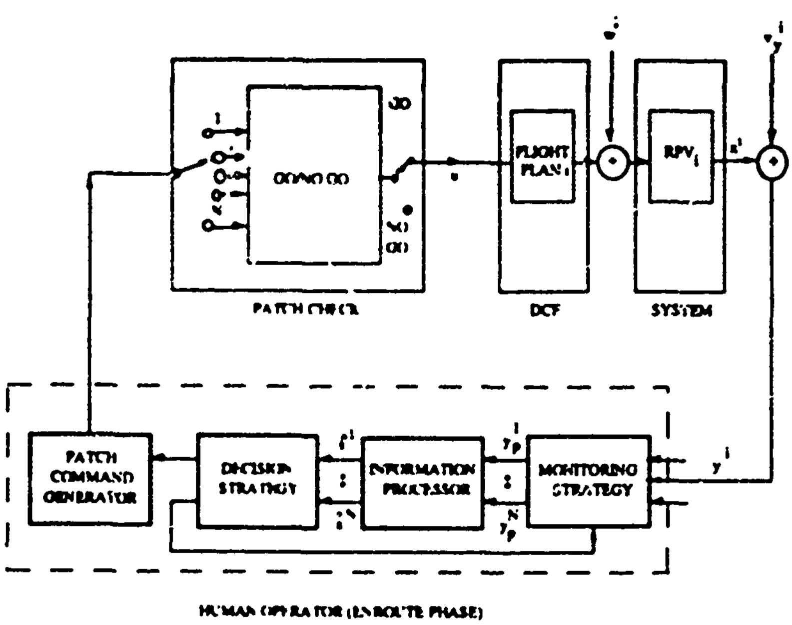


Fig. 4.8 DEMON model for RPV control task.

as communication among the operators can be considered. A block diagram modelling the flow of information and the control and decisions encountered by the human operator (enroute phase) is shown in Figure 4.8 (Muralidharan and Baron, 1980).

502. Path deviations arise from navigation errors and disturbances. Display information is assumed to be updated at discrete times. It is also assumed that arrival times and lateral deviation errors are presented separately and that only a single RPV can be selected for observation at a given time. Prior to frame update, then, the RPV enroute operator with N RPVs under control must also decide which among 2N+1 displays to monitor; the additional display is included to account for secondary tasks. The information processor in Figure 4.8 contains N Kalman filters to estimate the system state (i.e., the states of each RPV) and the uncertainty in that estimate. With this information, the processor can also compute the subjective probability of exceeding various error tolerances or the proximity to a waypoint (or desired goal). Thus, the information processor provides an assessment of the situation or a mental image of the traffic picture. The decision strategy generates both the monitoring and control choices. The operator's choices are assumed to be rational decisions governed by his knowledge of the situation, his goals and priorities and his instructions. These factors are incorporated in expected net gain criteria for monitoring and control. The expected net gain from a particular action is obtained by subtracting the cost of that action from its expected gain. The expected gain is the difference between the cost of events when no action is taken and the expected cost of events that may arise after this action.

503. In summary, the enroute operator's task is to monitor the trajectories and the estimated times of arrival (ETAs) of N vehicles, to decide if the lateral deviation or ETA error of any of these exceeds some threshold, and to correct the paths of those that deviate excessively by issuing acceptable patches. The drone control facility (DCF) contains the stored flight plans that drive the N subsystems RPV_i, i=1,2,...,N. They are usually "optimal" with respect to current terrain and other information. The system is constituted by the N RPVs undergoing monitoring/control. A simple non-linear representation of their dynamic behaviour was assumed for this analysis. Linearisation may be carried out if necessary for implementation of the model. The true status x^i of the i-th RPV may be different from the stored flight plans due to "disturbances" w^i. The reported status y^i will be different from the true status x^i due to reporting error v^i_y. The observed status y^i_p will depend on the reported status y^i and on the "monitoring strategy" (to be discussed later). The disturbances w^i and reporting error v^i_y were modelled by suitable random processes. The y^i are the displayed variables corresponding to RPV_i.

504. The monitoring strategy is needed, since the human must decide which RPV or which display to look at. This is important because his estimates of the true status of each RPV (and hence his patch decision strategy) will depend upon his monitoring strategy. The information processor models the processing that goes on in the human operator to produce the current estimate of the true RPV status from past observed status. This block is the well known control-theoretical model consisting of a Kalman filter-predictor. It produces the least-squares estimate of the true status and also the variance of the error in that estimate. The decision strategy models the process of deciding which, if any, RPV to patch. It considered the decision process to be discrete (it takes 5 sec to get a new display). The cost of making a patch would reflect the lost opportunity to monitor and/or patch other RPVs as well as

breaking radio-silence; the gain (negative cost) is the presumed reduction in error for the "patched" vehicle. The decision strategy attempts to minimise the (expected) cost.

505. The patch command generator generates the commanded patch. The developers investigated a strategy based on minimising a weighted sum of the time to return to the desired path and the total mean-square tracking error. The allowable paths were constrained by the RPV turning radius limits. The patch check performs a GO/NO GO check on the patch using conditions on turning radius, command link status, etc. The vector u denotes the patch control input to the RPVs. When there is no disturbance w^i and no patch control u then the N RPV subsystems follow the flight plan. A patching decision consists of deciding if the monitored RPV subsystem is to be patched. At most one of the RPVs may be patched at a given time. One idea of patching is to reduce deviations from the flight plan to below some threshold values. Once a decision is made to patch a particular RPV-subsystem, it is necessary to compute and execute the patch control. The purpose of a patch control is to guide the aircraft from its initial location and heading to intercept and fly along the planned flight patch. Various criteria may be considered to compute the optimal patch control, for example, a strategy that minimises the time to return to the planned flight path.

506. Operator limitations of observation noise, time delay and response bandwidth could be neglected in DEMON because they were insignificant in comparison to sensor measurement noise, display update rates and the time required for the discrete control inputs. However, a parameter was introduced into the information processing structure to account for the rate at which the operator's uncertainty about the vehicle state grows with time in the absence of further observation. This parameter relates to the operator's expectations concerning the disturbances perturbing the path of the RPV (as opposed to the true disturbances), as might be determined from instructions or through training. The mission goals were incorporated in the model in expressions for the expected net gains for monitoring and control. They included factors such as thresholds for allowable errors and costs for monitoring and patching control.

<u>When Used</u>

507. DEMON can be used in predictive studies and iteratively throughout design development.

<u>Procedures for Use</u>

508. The scenario-like DEMON approach is available in form of parametrised programme modules.

<u>Advantages</u>

509. According to Baron (1984), the DEMON model extended previous control-based approaches to a multi-task environment involving essentially discrete control decisions. For DEMON, control of each RPV represented a separate task, each with a payoff for maintaining errors within tolerance and for timely pop-ups and hand-offs. Inasmuch as only one RPV could be observed at any time, the DEMON operator had to rely on memory and prediction to decide when to monitor or serve a particular RPV and control objective. Another important advance was the introduction of the expected net gain decision making algorithm. This algorithm can be related to classical subjective expected utility criteria and other methods. It

also provides a very general approach to developing decision making criteria in multi-task situations.

Limitations

510. The main shortcoming is that the model does not include the procedural activities of the human operators.

Application Examples

511. Model results were obtained illustrating sensitivity of the performance predicted by DEMON to changes in parameters of the system (the number of RPVs to be controlled and the magnitudes of navigation and reporting errors) and in those describing operator behaviour. These results show that the model behaves reasonably, that the model parameters do significantly affect performance and that the monitoring and patching trends are as expected. For example, results from DEMON predicting how performance in controlling the RPVs varies with the number of RPVs under control, suggest that four vehicles can be maintained within tolerance by a single operator. Errors can exceed tolerances for five or six vehicles and a critical point exists around seven RPVs. This prediction is consistent with findings in the literature concerning the abilities of air traffic controllers.

Technical Details

512. The DEMON model is an example of the so-called top-down or analytic approach to human performance modelling. Such an approach begins with a mathematical characterisation of the task including the overall goals and the criteria for good performance. Then, one attempts to develop the assumptions about the human operator and the system that are necessary and sufficient to characterise performance in relation to the parameters of interest to system designers. The report of Muralidharan, Baron, and Feehrer (1979) includes complete technical documentation.

References

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In: W.B. Rouse (Ed.), Advances in man-machine systems (Vol. 1). Greenwich, CT: JAI Press.

Muralidharan, R. & Baron, S. (1979). DEMON: A human operator model for decision-making, monitoring, and control. Proc., 15th Annual Conf. Manual Control. Springfield, VA: National Technical Information Service (NTIS).

Muralidharan, R., Baron S.,& Feehrer C.E. (1979). Decision, monitoring, and control model of the human operator applied to an RPV control problem. (BBN Report No. 4075). Cambridge, MA: BBN Laboratories.

Muralidharan, R. & Baron, S. (1980). DEMON: A human operator model for decision making, monitoring, and control. J. Cybern. and Inform. Science, 3, 97-122.

Pew, R.W. & Baron, S. (1983). Perspectives on human performance modelling. Automatica, 19, (6), 663-676.

Future Needs

513. The supervisory control model PROCRU (section 4.3.6) is an extension of the DEMON approach.

4.3.6 A Model of Human Supervisory Control of Dynamic Systems (PROCRU)

Summary Description

514. The supervisory control model PROCRU (procedure-oriented crew model) has been developed by Baron et al. (1980, 1981) for analysing flight crew procedures. PROCRU incorporates both "by the book" procedures and more unconstrained control and monitoring behaviours. It models continuous tasks directly and also accounts for the effects of discrete control tasks and for the time to perform them. Thus, the model combines some of the features of psychologically-oriented models with those of control-theoretic models. PROCRU can be extended to a full range multi-operator model.

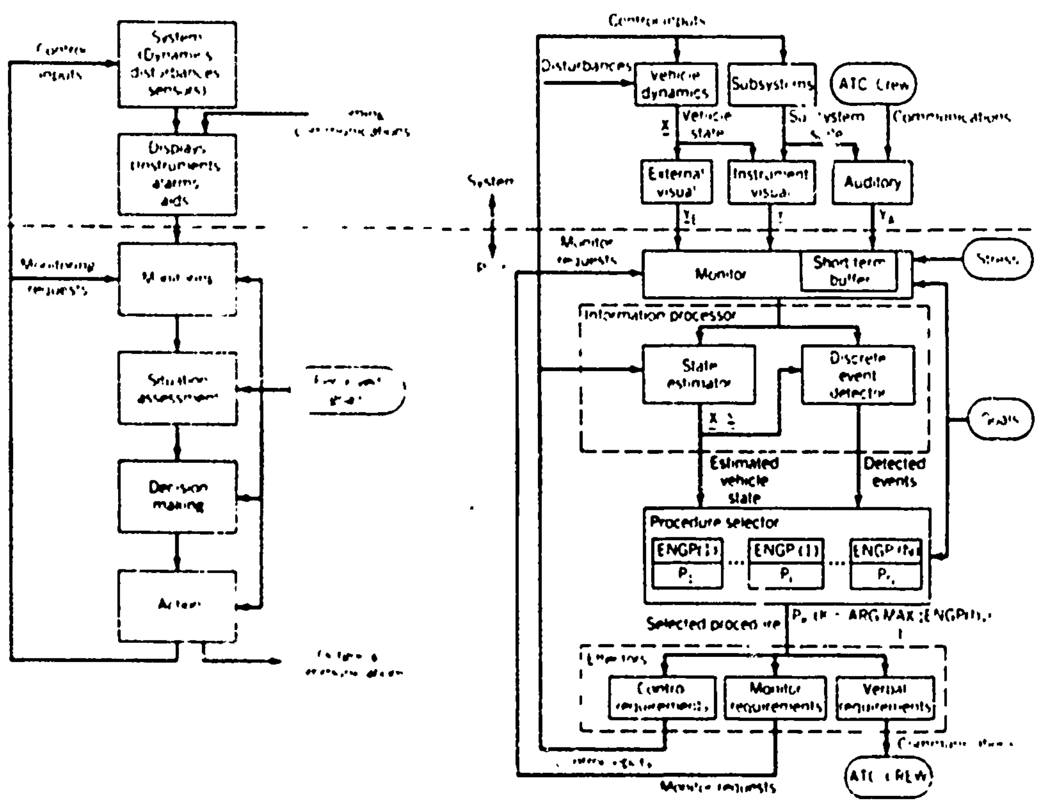


Fig. 4.9 Supervisory control model PROCRU.

History and Source

515. The monitoring and information processing portions of PROCRU are not unlike those of the optimal control model (OCM) or other models, though they have some novel features and extensions. The approach used in PROCRU has also been employed to develop

a supervisory control model called AAACRU for the commander/ gunner crew of an anti-aircraft artillery system (Zacharias et al., 1981). The general structure of this model is essentially identical to the PROCRU model structure, except that an explicit situation assessor has been added to the information processing function of the model. There is interest in developing more complex and complete supervisory control models based largely on a control-theoretic viewpoint. Presently there exists a conceptual approach to modelling operators of nuclear power plants (Baron et al., 1982). This model shares many features of the models described above.

Product and Purpose

516. The basic structure of the PROCU model is shown in Figure 4.9. The model is a closed-loop, man-machine simulation that incorporates elements that can represent a range of operator behaviours: cognitive and psychomotor, and continuous and discrete. Below, the major elements are discussed with special emphasis on the information processor, because of its central importance to the cognitive aspects of supervisory control and because it is in this part of the model that new extensions are suggested.

517. The supervisory control model PROCRU (Figure 4.9) builds on a series of operator functions and processes:

(1) Monitoring displays,

(2) Situation assessment,

(3) Decision to act - or not to act - based on that assessment, and

(4) Action to implement the decision.

518. These functions are implemented by various processors:

(1) A display processor selects an appropriate displayed quantity and accounts for sensory/processing limitations in observation.

(2) An information processor includes a mental model of the plant from which is derived a predict/correct logic for state estimation and prediction.

(3) A situation assessor provides a template matching scheme which checks symptoms against a template which is part of a procedure.

(4) A procedure selector includes major decision making at several levels. Choices are made on the basis of utility theory.

(5) A procedure effecter permits three types of actions: control, observation and communication. Time is associated with each action.

519. The lower half of Figure 4.9 illustrates the model for the human operator. The display processor portion of the model has two functions: it implements the conscious

observation decisions of the operator by selecting the appropriate displayed quantity; and, it accounts for sensory and processing limitations associated with observation. It is assumed that the operator is a single-channel processor of information. The selection of a particular source of information is governed by goal-oriented processes: thus, it will depend on the purposes for which information is being gathered. Much of the time, the choice of a particular source will be procedurally driven; i. e., the execution of specific operating procedures will direct the choice. However, for an activity that is not governed by well-defined procedures, the display selection process incorporated in the model will be designed to support that activity.

520. The information processor is comprised of three elements: an estimator/predictor, an event detector, and a situation assessor. The estimator/predictor performs two functions. First, it processes the observed information to update its estimate of the process state variables. And, second, it predicts the future evolution of the system on the basis of the estimate of the current state, known inputs to the process, and a mental model of the plant. The outputs of the estimator/predictor are the operator's a priori (before observation) and a posteriori (after update) estimates of the process state, the respective subjective estimation error covariances which, in general, will now differ from the true error covariances, and the residuals and their covariances. The discrete event detector is intended to model those aspects of operator information processing, other than estimation and prediction, of the process state variables. Typically, it will be concerned with determining or detecting that an event has occurred that helps to define a situation, thus enabling a procedure selection and execution. The event may be a transient, a request for action, an alarmed condition, or the verification of the accomplishment of an intended action. The inputs of the event detector are visual alarms, auditory information, the outputs of the state estimator/predictor, and the list in memory of possible events.

521. The situation assessor block of the supervisory control model is aimed at computing the probability or likelihood of a postulated situation. Its inputs are the outputs of the information processor and an ordered list of possible situations that are stored in memory. The ordering of the situations is assumed to be based on prior estimates of the probability of occurrence and on the potential consequences associated with the situation. As noted earlier, the situation may correspond to a single condition on a process variable or on equipment status. In such a case, situation assessment and event detection may coincide and the process is straightforward. More generally, a situation will be defined in relation to a larger set of conditions. However, in practice, the operator's assessment of the situation may not incorporate all the diagnostic elements.

522. The procedures are the means by which the operators organise and carry out their monitoring, situation assessment and control responses so as to accomplish their objectives. Procedures exist in manuals or they reside in memory, having been learned through training and/or experience. Both highly structured formal procedures and other less structured, but goal directed, responses are allowed. A formal procedure is a specific sequence of tasks or actions together with the situations that trigger those actions. Major decision-making at several levels takes place in the procedure selector block. The first decision is concerned with situation assessment and will involve some form of sequential test that will be performed to decide whether to accept/reject or defer the identification of a situation as discussed above. If the decision has been made to defer the identification of a situation, the modelled operator will have two options: collect more data using the current diagnostic test procedure; or, try a new diagnostic algorithm. Thus, a second decision is required.

When Used

523. PROCRU may be used for analysing flight crew procedures in commercial ILS approach to landing.

Procedures for Use

524. It is necessary to include a mathematical description of the system/environment in the supervisory control simulation model. The degree of detail and the level of complexity of the system model will depend on the specifics of the issues to be addressed. However, in general, the model must include those factors that are needed to perform a closed-loop system analysis, e. g., a state-variable description of the processes including any automatic control and engineered safety features, a description of potential disturbances, and a description of the instrumentation and display information provided for the crew (including information content, alarm set points, instrument or sensor noise, update rates and failure nodes).

Advantages

525. PROCRU represents one of the first major attempts to combine the monitoring and continuous control aspects of the OCM with a procedure selection and discrete control model.

Limitations

526. Some portions of the model are conceptual at this time in that some of its features have yet to be implemented (Baron,1984).

Application Examples

527. Only experimental applications have been reported.

Technical Details

528. Not available.

References

Baron, S., Muralidharan, R., Lancraft, R., & Zacharias G. L., (1980). PROCRU: A model for analyzing crew procedures in approach to landing. (BBN Report 4374). Cambridge, MA: BBN Laboratories.

Baron, S., Zacharias,G L. Muralidharan, R. & Lancraft, R., (1981). PROCRU: A model for analyzing flight crew procedures in approach to landing. Proceedings, 8th IFAC World Congress. Oxford: Pergamon Press.

Baron, S., Feehrer, C., Muralidharan, R., Pew, R. W., & Horwitz, P., (1982). An approach to modelling supervisory control of a nuclear power plant. (BBN Report 4811). Cambridge, MA: BBN Laboratories.

Baron, S. (1989). Engineering-based approaches to human performance modelling. In: McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Baron, S. (1984). A control theoretic approach to modelling human supervisory control of dynamic systems. In: W.B. Rouse (Ed.), Advances in man-machine systems (Vol. 1). Greenwich, CT: JAI Press.

Pew, R.W. & Baron, S. (1983). Perspectives on human performance modelling. Automatica, 19, (6), 663-676.

Sheridan, T. B. (1987). Supervisory control. In: G. Salvendy (Ed.), Handbook of human factors. New York: J. Wiley.

Zacharias, G.L., Baron, S., & Muralidharan, R., (1981). A supervisory control model of the AAA crew. (BBN Report 4802). Cambridge, MA: BBN Laboratories.

Future Needs

529. Not available.

4.3.7 The Human Operator Simulator (HOS)

Summary Description

530. The Human Operator Simulator (HOS) is a coordinated set of computer programmes which allows for the simulation of a total man-machine system performing a complex mission (Wherry, 1969; 1976; Lane et al., 1981; Meister, 1985). Thus, HOS simulates not only the behaviour of the operator, but the operating characteristics of the system hardware and software together with any sensors, targets, or other external data sources. It is intrinsically a bottom-up model in the sense that it begins with behavioural components and principles (i.e., micromodels), such as movement, information input, perception, and memory and systematically builds to a model that can perform task-oriented behaviour. Presently HOS-IV is available (Harris et al., 1987, 1988) and the concept of HOS-V has been formulated (Glenn, 1988). To build a simulation, inputs to the model typically include descriptions of the system design, procedures for using the system, human operator characteristics, and a mission scenario. A set of operator micromodels are available to the HOS user to assist in the development of the simulation. These micromodels contain algorithms, based on experimental literature, that can predict the timing and accuracy of basic human cognitive, perceptual, and psychomotor actions. The text of this HOS model summary is mainly based on Harris et al. (1987, 1988).

History and Source

531. Twenty years have elapsed since the original conceptual design specifications were set forth by Wherry (1969) for a generalised, goal-oriented, dynamically adaptive HOS computer programme. During this period HOS has been developed, applied, and modified on mainframe computers (Wherry, 1976; Lane et al. 1981). Recently, HOS has been restructured and revised to produce a 4th-generation version (HOS-IV). HOS-IV is already slated to be incorporated into the Automated Job Analysis Tool (AJAT). AJAT is a general purpose design and evaluation tool for man-machine systems with the objective of predicting the impact of individual differences in cognitive and psychomotor performance on total system performance (Glenn, Dick, and Bittner, 1987).

Product and Purpose

532. The Human Operator Simulator (currently HOS-IV) is a crewstation design evaluation tool which has unique features that distinguish it from other design tools. A detailed description of these and other features of HOS-IV can be found in the User's Guide (Harris et al., 1988). Unique features include:

(1) Capability to predict system performance by dynamic interactive simulation of the environment, the hardware/software system, and the operator.

(2) Library of human performance micromodels, based on experimental literature, to predict human performance times and errors.

(3) Simulation of an operator performing tasks and behavioural processes sequentially or in parallel.

(4) Flexibility to easily add to or modify the human performance micromodel library.

533. The central models resident in HOS are the operator models which will initially generate performance time and an indication of success/failure. The current version of HOS-IV contains cognitive (recall, attention, mental computations), perceptual (visual and auditory), and psychomotor (anatomy movement) models which are based on experimental data from the human performance literature. The HOS-IV outputs include:

(1) A timeline of events for the operator, system, and environment,

(2) User-defined measures of effectiveness, and

(3) Standard analyses, such as:
 (a) Mean time to complete an action,
 (b) Number of times an action is performed,
 (c) Proportion of the operator's time spent on each action, and
 (d) Error analysis by action.

534. The revised modelling capability in HOS-IV incorporates major improvements in the areas of efficiency, usability, and adaptability. HOS-IV, in particular, provides ready access to a library of standard models for items commonly included in simulations such as controls, displays, etc. In addition, the HOS analyst will be able to tailor the models to the needs of the particular application by specifying the appropriate level of detail for each model (or, if desired, incorporating his or her own models). All models included in HOS-IV are written utilising the same HPL language as used to define simulation actions. The human performance micromodels of HOS-IV are:

(1) Psychomotor Micromodels
 (a) Eye Movement
 (b) Hand Movement
 (c) Control Manipulation
 (d) Handprinting
 (e) Walking

(2) Cognitive-Perceptual Micromodels
 (a) Visual Perception
 (b) Short-Term Memory Store
 (c) Short-Term Memory Retrieve
 (d) Decision Making

(3) Communication Micromodels
 (a) Listening
 (b) Speaking

(4) Fatigue Micromodel
 (a) Modulating Effects on Human Performance

(5) Planned Micromodel Development
(a) Tracking
(b) Target Search
(c) Task Switching

535. The eye movement model determines the time required to move the eye from one fixation point to the next. The hand movement action determines the time required to move the hand to a new position. The movement time model is based on Fitt's law which states that movement time is a linear function of the information content or difficulty of movement. The control manipulation model determines the simulation time charge for manipulating four different types of controls - pushbutton, toggle, rotary dial, or trackball. The handprinting model determines a time charge for printing a specific number of characters as specified by the user. The walking model determines the time required per foot of travel.

536. The visual perception model determines the time required to perceive a visual target assuming the eye is already at the fixation point and that the target is clearly visible. The short term memory model consists of two parts - memory storage and memory retrieval. The memory storage component simulates entry of new items into a push-down memory stack. The memory stack has a maximum size of seven items. The memory retrieval component simulates memory decay as well as retrieval search times.

537. The listening model determines the time to listen as well as the probability of a listening error based on the signal-to-noise interruption frequency. The speaking model determines the time to read aloud a user-specified number of words. The speaking time-per-word varies depending on whether the user specifies that the words are elements of a small or a large vocabulary.

538. The modulating effects of mental fatigue on human performance are represented by a separate model that is based on numerous experimental studies using critical flicker fusion (CFF) as measuring methodology. Specifically, the model estimates the percentage decrement in cognitive performances for a given time t.

539. Three performance models are planned for augmentation of the current set in HOS-IV. The planned set of tracking and continuous manual control models is essentially based on a composite linkage of the already existing psychomotor and cognitive micromodels. Tracking capabilities (e.g., compensatory, pursuit, and learning) have been demonstrated in previous studies (Glenn, 1982; Lane et al., 1981) The planned set of target search models will be based on a composite linkage of the cognitive-perceptual micromodels. Visual search capabilities have been demonstrated already with earlier HOS applications (Lane et al., 1981). The implementation is intended for allowing various search strategies. The planned task switching model is based on evidence of individual-difference-related and other temporal cost for mental switching between tasks

540. The micromodel modes of operation have been considerably extended. HOS-IV permits the user to describe operator behaviours so that the associated micromodels occur either sequentially or in parallel as appropriate for the particular tasks to be simulated. For sequential operations, each micromodel, or elemental behaviour within the task, must be completed before any other elemental steps or other operator tasks may occur. Parallel operations simulate situations where the operator is performing tasks concurrently. This structure allows parallel execution of non-competing behavioural processes - for example, a

cognitive process will not interfere with a motor process. In general, micromodels within a particular grouping - cognitive/perceptual, communication, or psychomotor - compete with each other and can stalemate performance. Previous versions of HOS had been restricted to the assumption that the human operator can only do one thing at a time (single-channel processing assumption).

When Used

541. HOS is particularly well suited to performing analyses in which the operator's workstation layout is the main concern and where performance degradation results primarily from work congestion in reading displays and manipulating controls. In addition, HOS seems to be suited for educational and instructional purposes, e.g., helping to introduce modelling and and simulation of human-machine systems

Procedures for Use

542. HOS is a general purpose simulation tool for modelling man-machine systems. The required inputs to the model are descriptions of the system design, procedures for using the system, human operator characteristics, and a mission scenario. A set of operator micromodels are available to the HOS user to assist in the development of the simulation. HOS provides a capability for a comprehensive human-system simulation. Figure 4.10 presents the components developed for an application. During a typical implementation, such as for a model of a radar operator and crewstation, the analyst first determines the allocation of functions between the human operator and the machine. The analyst then describes the environment (e.g., number, location, speed, and bearing of enemy targets); the hardware system (e.g., radar sensor and signal processors, displays, and controls); and the operator procedures and tactics for interacting with the system and for accomplishing mission goals.

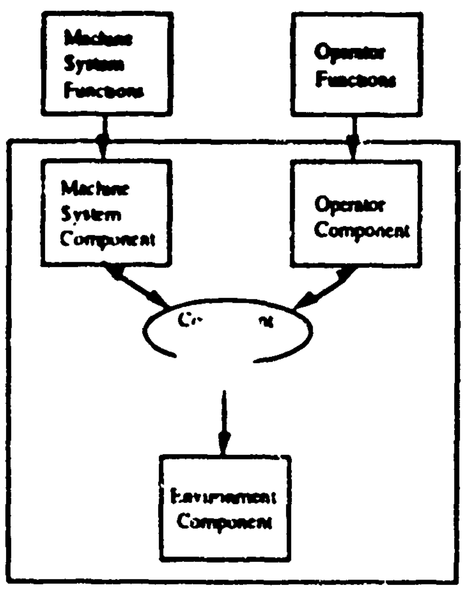


Fig. 4.10 Application of the Human Operator Simulator (HOS).

543. Interface descriptions between the operator, system, and environment are also developed to suit the needs of a particular application. In a radar system simulation, a hardware-environment interface routine would determine which enemy targets were within the radar detection range at any given temporal snapshot. An operator-environment interface routine could determine the effects of heat, cold, drugs, or other stressors on human performance timing and accuracy. The operator-machine interface models could establish the time and accuracy of an operator performing such tasks as reading alphanumeric information from displays, manipulating controls, searching for targets in a particular field-of-view, or physically moving objects from one location to another.

Advantages

544. One primary feature of HOS, compared to other simulations, is that HOS is now rule-based, incorporating current artificial intelligence techniques to structure the simulation. An input requirement to HOS is a set of rules which activate a set of actions during the course of the simulation only when the conditions are appropriate. Defining these rules facilitates a top-down approach to the design of a simulation since it allows the user to design the simulation flow independent of the implementation of low level simulation actions and models. It also allows the simulation to more closely mimic reality since operators usually make decisions based on an implicit or explicit set of rules when responding to a particular situation.

Limitations

545. The level of detail required by HOS and the restrictions on multiple operators and communications may make HOS inappropriate for large systems involving complex interactions among more than one operator.

Application Examples

546. One area of model application has been a series of part-task simulations. Tasks in these studies closely resembled operator functions in systems but were less complex to allow more precise structuring of task conditions. The objectives were to verify the additivity of times generated by the micromodels and to examine simulation performance in situations more closely approximating real systems than laboratory experiments. These simulations included (Lane et al., 1981):

(1) A divided attention study in which the operator performed a manual tracking task with interference from a secondary task whose frequency and duration were varied.

(2) A mail-sorting simulation which required the operator to use simple decision rules combined with keyboard entry.

(3) A subset of the LAMPS helicopter Air Tactical Officer (ATO) functions, combining CRT tracking and control manipulation with key entry.

547. As the use of HOS has matured, simulations have been developed for several large-scale systems. Results of these applications are another type of HOS validation. The major simulations conducted with HOS to date are (Lane et al., 1981):

(1) A simulation of the Air Tactical Officer (ATO) on-board the U. S. Navy's LAMPS helicopter during a generalised air surveillance mission.

(2) A simulation of the Sensor Station 3 (non-acoustic) operator (SS-3) on the U. S. Navy's P-3C ASW patrol aircraft during a reconnaissance mission.

(3) A simulation of the P-3C Sensor Station 1 (acoustic) operator (SS 1) during an open ocean convoy escort mission.

(4) A simulation of the pilot on NASA's Terminal Configured Vehicle (TCV) during the approach and landing phases of both a curved and straight landing under both manual and automatic control.

Technical Details

548. Currently HOS-IV is implemented in Microsoft C (Version 4.0) on an IBM PC-AT with enhanced graphics, a mouse, additional random access memory, and auxiliary storage devices. HOS-IV incorporates many new features with the most notable being (1) a new knowledge representation scheme, (2) a user-oriented interface, and (3) enhanced modelling capabilities. The primary design goal of the HOS-IV development was to provide for effective application simulations without placing an excessive burden on the human analyst. The HOS simulation facility is also highly transportable, allowing implementation on IBM compatible microcomputers. Concepts for the development of HOS-V have been formulated (Glenn, 1988).

References

Glenn, F. A. (1982). A discrete learning model for manual pursuit tracking. Proc., IEEE Conf. on Cybernetics and Society. New York: IEEE.

Glenn, F. A., Dick, A.O., & Bittner, A.C. (1987). Prediction of personnel requirements for system operation. Proc., Human Factors Society, 32nd Annual Meeting. Santa Monica, CA: Human Factors Society.

Glenn, F. A. (1988). A human operator model management environment. Proc., Human Factors Society, 32nd Annual Meeting. Santa Monica, CA: Human Factors Society.

Glenn, F. A. (1988). Development of a human operator simulator version V (HOS-V): Concept formulation (TR 88 04 30). Blue Bell, PA: CHI Systems, Inc.

Harris, R. M., Iavecchia, H. P., Ross, L. V., & Shaffer, S. C. (1987). Microcomputer human operator simulator (HOS-IV). Proc., Human Factors Society, 31st Annual Meeting. Santa Monica, CA: Human Factors Society.

Harris, R. M., Bare, C., Ross, L., Scolaro, D., Wright, D., & Kaplan, J. (1988). Human operator simulator (HOS-IV), user's guide (Tech. Report 2072-B). Willow Grove, PA: Analytics, Inc.

Harris, R. M., Iavecchia, H. P., Bittner, A. C. (1988). Everything you always wanted to know about HOS micromodels but were afraid to ask. Proc., Human Factors Society, 32nd Meeting. Santa Monica, CA: Human Factors Society.

Harris. R. M., Iavecchia, H. P. & Dick, A. O. (1989). The human operator simulator (HOS-IV). In McMillan, G. R., Beevis, D., Salas, E., Strub, M. H., Sutton, R., & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum.

Lane, N. E., Strieb, M. I., Glenn, F. A., Wherry, R. J. (1981). The human operator simulator: An overview. In J. Moraal & K.-F. Kraiss (Eds.). Manned systems design: Methods, equipment, and applications. New York: Plenum Press.

Meister, D. (1985). Behavioral analysis and measurement methods. New York: J. Wiley.

Wherry, R. J. (1969). The development of sophisticated models of man-machine-systems. Proceedings, Symposium on Applied Models of Man-Machine Systems Performance Columbus, OH: North American Aviation.

Wherry, R. J. (1976). The human operator simulator - HOS. In T.B. Sheridan & G. Johannsen (Eds.). Monitoring behavior and supervisory control. New York: Plenum Press.

Future Needs

549. Models like HOS, with the ability to integrate a realistic operator component into a systems modelling framework, will become more critical as systems depart from traditional roles for the operator. Human factors engineering technology must keep pace with equipment technology, providing techniques which allow credible, objective and detailed statements about probable system (not just operator) performance. Without such techniques, the inappropriate use of operator capabilities cannot be prevented. Integrating HOS with general purpose simulation technology (e.g., SAINT, SLAM, etc; (see Chapter 8) could be helpful.

4.3.8 The Siegel-Wolf Model

Summary Description

550. According to Pew et al. (1977), the Siegel-Wolf model, which appeared in 1961, has been of continuing interest to human performance analysts and has been consequently extended (Siegel and Wolf, 1981). It embodies some of the early network concepts (e.g., PERT; Pew et al., 1977; Chubb et al., 1987) and depends significantly on relationships between required task times and available task times for measures of completion probability and stress. It is also built upon conceptual structures and empirical observations drawn from psychology and human factors and, as such, can be considered a model in its own right. The present and continuing interest in the Siegel-Wolf model mainly stems from successfully including model components of:

(1) human performance under stress,

(2) subtask execution time under stress,

(3) multi-operator performance, and

(4) their relationship to performance reliability and human error in man-machine systems.

551. The Siegel-Wolf models are representative of a type, for instance, stochastic or network, but more importantly they have been outstandingly successful - if success is defined as the variety of situations to which they have been applied. There are three such models:

(1) the 1- to 2-man model,

(2) the 4- to 20-man model, and

(3) the 20- to 99-man model.

552. Only the first is considered here. The purpose of the 1- to 2-man model is to serve as a tool for system designers during development, and to indicate where the system may over- or underload its operator. The model simulates maintenance or operator tasks simply by identifying personnel as operators or technicians and the tasks as operator or maintenance tasks. It predicts task completion time and the probability of successful task completion. It also seeks to determine whether or not an average operator will successfully complete required tasks, how success probability changes for various performance shaping factors, and the operator proficiency required by the system. It is interesting to note that the concept of performance moderator functions is used within present model developments too (Laughery and Gawron, 1984). The text of this Siegel-Wolf model summary is in parts based on Pew et al. (1977) and Meister (1985).

History and Source

553. The first and some later versions of the Siegel-Wolf model were developed for the US Office of Naval Research.

Product and Purpose

554. The methodology underlying the Siegel-Wolf model can be applied to any type of system or task. The basic assumption of the model is that operator loading is the basic element in effective man-machine system performance and that the variety of loading effects are compressed into one variable called stress. Stress may be caused by (1) falling behind in time on an assigned task sequence; (2) a realisation that the operator's partner is not performing adequately; (3) inability to complete successfully a subtask on the first attempt and the need to repeat the subtask; (4) the need to wait for equipment reactions. Basic mechanisms of the model can be understood from a description of its parameters:

(1) Execution time t_{ij} represents the amount of time required by operator j to complete subtask i.

(2) Completion probability P_{ij} represents the probability that subtask i is successfully performed by operator j.

(3) The parameter T_j, the mission time limit, specifies the total time available for operator j to perform a task that is constituted by the sequential subtasks i = 1, ..., n.

(4) The parameter F_j accounts for the variance among individuals operating the system. Thus the model is able to simulate operators who usually perform faster or slower than the average operator.

(5) The stress level S_{ij} of operator j is operationally defined as the ratio of (a) how much is left to do when performing subtask i to (b) the amount of time available in which to do it:

$$S_{ij} = \frac{\bar{T}^{E}_{ij}}{T_j - T^{u}_{ij}}, \tag{1}$$

where T_j is the total time available; T^{u}_{ij} is the time elapsed up to but not including accomplishment of subtask i; and $\bar{T}^{E}_{ij}$ is the average time required for completion of all remaining subtasks, assuming no failures.

555. The primary relationships of the model can be expressed in form of two submodels:

(1) Submodel of Performance under Stress. Task/subtask completion probability $\overline{P_{ij}}$ is a function of stress level S_{ij} and stress threshold M_j. This probability of success increases linearly with stress until it becomes unity at the stress threshold. At this value, it assumes the average (i.e., input value), $\overline{P_{ij}}$ and then decreases linearly until it reaches a constant value.

(2) Submodel of Execution Time under Stress. Task/subtask execution time t_{ij} is a function of stress level S_{ij} and stress threshold M_j. The average execution times are decreased with increasing stress until stress assumes the threshold value M_j; beyond M_j the average execution times are increased linearly with increasing stress.

556. The model outputs include a considerable amount of data for each operator. A run summary might contain total number of runs, number and percent of successful runs, average time used over N runs, average time over run, average waiting time, average peak and final stress, the number of times a subtask was failed or ignored, the time spent in repeating failed subtasks, and the average time that the subtask was completed.

557. An important aspect of Siegel-Wolf model is its suitability for use in simulations of the performance of multi-operator systems. In this role, the technique employs a number of self-contained models to determine expected task time as a function of group performance proficiency, overtime load, morale, number of persons, and nominal time. Multi-operator aspects of the model are described in Chapter 6.

When Used

558. The Siegel-Wolf model has been used in numerous studies (at least 10 validation studies have been reported; Meister, 1985) with respect to design analysis and prediction of system performance. Firstly, the model can supply an absolute estimate of the system reliability to be anticipated when the system becomes operational (e.g., the system will eventually perform with a reliability of .99). Secondly, the model can compare alternative system configurations to determine which should be selected for implementation or to determine redesign requirements for a system which cannot satisfy system requirements. Thirdly, if the simulation model can predict the system effectiveness of one configuration, it can also predict the effectiveness of another, and compare the two estimates.

Procedures for Use

559. To apply the model, one first performs an analysis and identifies which tasks are essential to completion of the mission and which are unessential. For each of these, the following data are specified, using the available sources:

(1) The average time required by the operator to perform each task.

(2) The average standard deviation about the average time.

(3) The average probability of successfully performing each task.

(4) An estimate of the extent to which successful performance of each task is required for completion of the total mission.

(5) The waiting time (if any). This represents the time elapsed between the start of the mission and the start of each task, during which no action can be taken by the operator.

(6) The next task to be performed given failure to accomplish a current task.

(7) The next task to be performed given successful accomplishment of a current task.

560. The level of data input is fairly molecular, describing individual discrete perceptual and motor actions. Sources of input data are varied. Data are collected from task analysis, formal experiments, informal measurements, simulator measurements, literature search or personal interviews. Much of the input data is gathered by direct questioning of expert operators; the data gathering process is relatively informal. Although the model makes use of data banks, such as they are, it is likely that some new input data must be gathered for each new application of the model.

561. The simulation operates through its Monte-Carlo sampling process to arrive directly at the end result. Success or failure of the entire task or mission is not dependent on the probability of accomplishment of any single subtask, but whether or not the operator completes all essential subtasks in the required time. Each individual subtask has an effect on ultimate system success but not necessarily a primary one. As a consequence, all the computer does at the end of a series of computer runs is to divide the number of successful runs by the total number of runs performed to arrive at an estimate of effectiveness. Because the simulation of any individual task is based in part on a random process, it is necessary to repeat the simulation a number of times to obtain sufficiently representative performance data for each set of conditions. A value of N, usually 100 to 200 iterations, is selected prior to the simulation.

Advantages

562. According to Pew et al. (1977), the model has a number of significant virtues. One is a mechanism for modifying performance in accord with stress. A second is the Monte-Carlo component of the technique. Rather than predicting a judgment of the performance of the system on the sum of single subtask expected values, the model expressly considers that operators will differ in their performance of the same subtask and that the same operator will exhibit differences in successive repetitions. This capacity to encompass both within-operator and between-operator variability is an important desideratum in performance modelling.

Limitations

563. Several weaknesses exist (Pew et al., 1977). There is, first of all, the question of whether valid structural representations and data can be assembled for a system in which possible interactions may be only dimly understood. Secondly, there is a question concerning the soundness of the assumption that the means and standard deviations specified are, in fact, the means and standard deviations of normal distributions. Thirdly, there is some concern over the ability of the model to yield accurate predictions of performance in situations where observable task density is low (as, for example, in tasks requiring a great deal of monitoring and signal detection but only occasional system input) or where the performance of several operators functioning in parallel must be assessed.

Application Examples

564. The model has been validated by Siegel, Wolf and their colleagues (see references) in the course of their simulations of a wide variety of unitary and dual operator tasks and seems to represent the observed relationships between stress and performance quite well. Major experimental emphasis has centered on the search for appropriate values of each operator's stress threshold for any given simulation. An introduction into the use of the model as well as sample applications are given by Siegel and Wolf (1969). Further applications include carrier-aircraft landings, missile launchings, and submarine operations.

Technical Details

565. A detailed model description is given by Siegel and Wolf (1969); an error in the mathematical specification has been mentioned and corrected by Pew et al. (1977). A state-of-the-art report of the model is given by Siegel and Wolf (1981).

References

Chubb, G. P., Laughery, K. R., & Pritsker, A. A. B. (1987). Simulating manned systems, In G. Salvendy (Ed.). Handbook of human factors. New York: J. Wiley.

Laughery, K. R. & Gawron J. M. (1984). Human factors moderator functions. (MOPADS, Techn. Report). Fort Bliss, TX: US Army Research Institute Field Unit.

Meister, D. (1985). Behavioral analysis and measurement methods. New York: J. Wiley.

Pew, R. W., Baron, S., Feehrer, C. E., & Miller, D. C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation. (BBN Report No. 3446). Cambridge, MA: BBN Laboratories.

Siegel, A. I. & Wolf, J. J. (1961). A technique for evaluating man-machine systems design. Human Factors, 3, 18-28.

Siegel, A. I. & Wolf, J. J. (1962). A model for digital simulation of two-operator man-machine systems. Ergonomics, 5, 557-572.

Siegel, A. I. & Wolf, J. J. (1969). Man-machine simulation models: Performance and psychosocial interaction. New York: J. Wiley.

Siegel, A. I., Wolf, J. J., & Cosentino, J. (1971). Digital simulation of the performance of intermediate size crews: Application and validation of a model for crew simulation. Wayne, PA: Applied Psychological Services.

Siegel, A. I., Wolf, J. J., & Cosentino, J. (1971). Application and validation of a model for crew simulation. Wayne, PA: Applied Psychological Services.

Siegel, A. I., Lautman, M. R., & Wolf, J. J. (1972). A multimethod-multitrait validation of a digital simulation model. Wayne, PA: Applied Psychological Services.

Siegel, A. I. & Wolf, J. J. (1981). Digital behavioral simulation - State-of-the-art and implications. Wayne, PA: Applied Psychological Services.

Siegel, A. I., Wolf, J. J., & Pilitsis, J. (1982). A new method for the scientific layout of workspace. Applied Ergonomics, 13, (2) 87-90.

Future Needs

566. It is likely that the model can be employed outside of the Siegel and Wolf simulation with a reasonable promise of success. Therefore user-friendly implementations of the Siegel-Wolf model should be made available (e.g., on the basis of SAINT, SLAM, MicroSAINT, or HOS; see Chapter 8).

4.4 INTRODUCTION: MODELS OF HUMAN ERROR

567. The two fundamental human performance parameters are speed and accuracy. Thus human accuracy, or its counterpart, human error, should be a major factor in the study of operator performance, and in the prediction of systems readiness, effectiveness, and reliability. Despite the demonstrated role of human error in recent nuclear power, commercial airline, and military accidents, human reliability (or human error) is seldom studied directly during systems development or design. Sometimes it is a consideration at the function allocation stage of analysis, but in qualitative terms only (e.g. the "man is not reliable, but does not suffer sudden failure" approach of the Fitts List). At the detailed design stage, it is implicit in the application of guidelines and standards for design of the man-machine interface and design for maintenance, many of which minimise the probability of human error. Human error probability is seldom used directly as a factor in design tradeoffs, however.

568. Reliability engineers have an arsenal of empirical and mathematical techniques for estimating the reliability of equipment, and some human factors specialists have seen a need for a corresponding treatment of human reliability. The hope is not only that human behaviour can be expressed in the same reliability terms as equipment, but that measures of human reliability can be combined with measures of equipment reliability to give an estimate of the reliability of the system as a whole. The various techniques and approaches that have been offered for dealing with one or more aspects of this problem have gradually formed the area of human reliability analysis (HRA).

569. In an early review of methods of predicting human error, Meister (1964) reported the fault-tree approach of THERP (Technique for Human Error Rate Prediction), and a Monte-Carlo simulation approach by Miller. Subsequent reviews by Meister (1973, 1984) suggest that those two approaches remain the mainstream of efforts to model human error behaviour. This review supports that viewpoint. Although Miller's approach does not appear to be widely used, there is growing use of Monte-Carlo simulation, as exemplified by the Siegel-Wolf model described in section 4.3.8 of this report.

570. Of the two mainstream approaches, the Siegel-Wolf model appears the most relevant to the preliminary stage of analysis and design. THERP appears to be more suited to later stages of design, when specific details of the man-machine interface have been decided. As Swain and Guttman (1975) indicate, THERP has also been used to evaluate operational systems and procedures.

4.5 OVERVIEW AND RECOMMENDED REFERENCES: MODELS OF HUMAN ERROR

Definition of Human Error

571. Following the definition of Rigby (cited by Miller and Swain, 1987), human error is any member of a set of human actions that exceeds some limit of acceptability. It is an out-of-tolerance action, where the limits of acceptable performance are defined by the system. Most errors are unintentional, inadvertent actions that are inappropriate in the given situation. There are some errors that are intentional. They occur when someone intends to perform an act that is incorrect, but believes it to be correct or to represent a better method. Malevolent behaviour, on the other hand, is not considered to be part of human error in this treatment. Not all human errors result in system degradation. An error can be recovered or corrected before it results in undesirable consequences to the system.

572. Human performance is adaptable, variable, non-linear, and includes monitoring of

"own" performance. This results in error rates which are low (on the order of tenths of a percent) and irregular. Most laboratory experiments which attempt to study error place the human subjects in conditions which result in unrealistically high error rates, in order that sufficient data can be gathered in a practical time period. Consequently much of the study of actual operator error tends to be based on case studies. Swain and Guttman (1975) have argued that this results in any test of the validity of such models being unidirectional: case studies can only indicate that the observed reliability is not less than the predicted reliability.

Definition of Human Reliability

573. Reliability is the antithesis of error likelihood. It is the probability that no errors occur. Reliability is conventionally defined as the probability of successful performance of a mission. Meister (1966) defined human reliability as

> "The probability that a job or task will successfully be completed by personnel at any required stage in system operation within a required minimum of time (if the time requirement exists)."

Swain and Guttmann (1983) defined human reliability as

> "The probability that a person (1) correctly performs some system-required activity in a required time period (if time is a limiting factor) and (2) performs no extraneous activity that can degrade the system."

Human Error Probability

574. The basic expression of error likelihood is the human error probability (HEP). The HEP is the probability that when a given task is performed, an error will occur. An HEP is calculated as the ratio of errors committed to the number of opportunities for that error, or an estimate of that ratio.

$$HEP = \frac{\text{number of errors}}{\text{number of opportunities for error}} \qquad (1)$$

575. The denominator represents the exposure to the task or task element of interest and is often difficult to determine because the opportunities may be covert, unrecorded, or part of a procedure whose steps appear to be continuous. The assumed frequency distribution of the HEP variable over people and conditions has not been consistent in the literature. For instance, Askren and Regulinski (1969) found that the Weibull distribution gave the best fit to their data. Swain and Guttmann (1983) have assumed a log-normal distribution. When the performance of skilled people is considered, it is reasonable to assume that most HEPs fall near the low end of the error distribution.

Performance-Shaping Factors

576. The concept of performance-shaping factors (PSFs) has been introduced by Swain (see Miller and Swain, 1987) to describe any factor that influences human performance. External performance-shaping factors are those that predispose human operators to increased errors.

Examples are given by Miller and Swain (1987):

(1) Inadequate work space and work layout.

(2) Poor environmental conditions.

(3) Inadequate human engineering design.

(4) Inadequate training and operating procedures.

(5) Poor supervision.

577. Internal performance-shaping factors are human attributes such as skills, abilities, and attitudes that the operator brings to the task. If training has been adequate, however, internal PSFs generally have less impact than external PSFs on human reliability. Some examples of internal performance shaping factors are (Miller and Swain, 1987):

Training/experience	Emotional state	Physical condition
Skill level	Perceptual abilities	Sex differences
Intelligence	Task knowledge	Strength/endurance
Motivation/attitude	Social factors	Stress level

578. Stress, a very important internal PSF, is the body's physiological or psychological response to an external or internal stressor. Stress usually has a nonmonotonic effect on performance. At very low levels of stress, there is not enough arousal to keep a person sufficiently alert to do a good job. Similarly, at high stress levels, performance usually deteriorates as the stressor increases or persists for long periods of time. Somewhere between low and high levels of stress, there is a level associated with nearly constant performance, called the optimal level of stress. Disruptive stress can increase the possibility of error by a factor of 2 to 5 according to Swain and Guttmann (1983), and extremely high levels of stress can result in even higher degrees of performance degradation. Another internal performance-shaping factor that can have a significant influence on task performance and human error is experience. A combination of stress and inexperience can increase the error probability of a human operator by a factor of as much as a factor of 10. To account for the effects of stress and experience, Swain uses the following model to estimate the increase in human error probability by a factor (Miller and Swain, 1987):

Table 4.1 Effects of Stress and Skill Level on Error Probability

Stress Level	Skilled Operator	Novice Operator
Very Low	x2	x2
Optimum	x1	x1
Moderately High	x2	x4
Extremely High	x5	x10

Error Taxonomies

579. The prediction of human error within the systems design/development cycle would be facilitated by a human error taxonomy which relates the characteristics of an operator's task

to an error rate. Thus a large number of error taxonomies have been proposed (Miller and Swain, 1987; Norman, 1981; Rasmussen et al., 1987; Rouse, 1983; Swain and Guttmann, 1983). Unfortunately with the exception of the early work by Munger, Smith and Payne (1962), which is sometimes referenced, there is no widely accepted taxonomy of human error. The US National Research Council sponsored the development of a human error data base for use in the design of nuclear power plants (Miller and Swain 1987), but that data base is not widely referenced.

580. The two principal approaches to modelling human error, then, are the engineering reliability calculation approach, typified by THERP, and the Monte-Carlo simulation approach typified by the Siegel-Wolf model. Although other approaches have been, or are being, developed, they are not in widespread use. For example Thomas (1978) reviews several models, one of which uses a Markov chain approach to modelling, but none appear to be referenced widely. Dhillon (1986) briefly reviews six methods, two of which are closely related to THERP, and three of which are not referenced widely. Several other models are reviewed by Embry (1976), Meister (1984), Miller and Swain (1987), and Pew, Feehrer, Baron, and Miller (1977).

4.5.1 References

Askren, W.B. & Regulinski, T.L. (1969). Quantifying human performance for reliability analysis of systems. Human factors, 11, (4), 393-396.

Embry, D. (1976). Human reliability in complex-systems. An overview. Warrington, UK: National Centre for Systems Reliability, United Kingdom Atomic Energy Authority.

Dhillon, B.S. (1986). Human reliability with human factors. New York: Pergamon Press.

Meister, D. (1964). Methods of predicting human reliability in man-machine systems. Human Factors, 6, 621-646.

Meister, D., & Rabideau, G.F. (1965). Human factors evaluation in system development. New York: Wiley & Sons.

Meister, D. (1966). Human factors in reliability. In W.G. Ironson (Ed.). Reliability handbook. New York: McGraw-Hill.

Meister, D. (1973). A critical review of human performance reliability predictive methods. IEEE Trans. on Reliability, 22, (3), 116-123.

Meister, D. (1984). Human reliability. In F.A. Muckler (Ed.). Human factors review. Santa Monica, CA: The Human Factors Society.

Meister, D. (1985). Behavioral analysis and measurement methods. New York: Wiley Interscience.

Miller, D.P. & Swain, A.D. (1987). Human error and human reliability. In G. Salvendy (Ed.) Handbook of human factors. (pp. 219-250). New York: Wiley & Sons.

Munger, S.J., Smith, R.W., & Payne, D.P. (1962). An index of electronic equipment operability: Data store. (AIR-C43-1/62-RP(1)). Pittsburgh. PA: American Institute for Research.

Norman, D.A. (1981). Categorization of action slips. Psychological Review, 88, (1), 1-15.

Pew, R.W., Baron, S., Feehrer, C.E. & Miller, D.C. (1977). Critical review and analysis of performance models applicable to systems evaluation. (BBN Report 3446). Cambridge, MA: Bolt, Beraneck, and Newman.

Rasmussen, J., Duncan, K., & Leplat, J. (1987). New technology and human error. Chichester: John Wiley & Sons.

Rouse, W.B. & Rouse, S.H. (1983). Analysis and classification of human error. IEEE Trans. on syst., man, and cybern., 13, (4), 539-549.

Swain, A.D., and Guttman, H.E. (1975). Human reliability analysis applied to nuclear power. Proceedings, Annual Reliability and Maintainability Symposium. (pp. 110-115).

Swain, A.D. & Guttmann, H.E. (1983). Handbook of reliability analysis with emphasis on nuclear power plant applications. Washington, DC: U.S. Nuclear regulatory commission.

Thomas, M.U. (1978). Some models of human error for man-machine system evaluation. (Tech Report No. 75-9), Detroit, Michigan: Chrysler Corp. and University of Michigan, Dept. of Industrial and Operations Engineering.

4.6 MODEL SUMMARIES: MODELS OF HUMAN ERROR

4.6.1 Technique for Human Error Rate Prediction (THERP)

Summary Description

581. THERP is a technique for predicting the impact of human operator errors on system operation. The technique is similar to conventional reliability engineering approaches, with modifications to cater to the greater variability, unpredictability and interdependence of human performance. THERP permits the estimation of the probability that an operation will result in an error, and the probability that an error, or class of errors will result in system failure. The basis of THERP is the preparation of a Human Reliability Analysis (HRA) event tree, to which are attached probabilities of success or failure of each task. The HRA event tree is a key feature of the technique which distinguishes it from the engineering technique of fault tree analysis. HRA event trees reflect a sequence of operator activities, working forward in time. Fault trees start with a possible fault, and work backwards to identify contributory events.

582. The HRA event tree shows the sequence of tasks with branches at each point where the operator could make an error (Figure 4.11). The next stage of analysis involves the assignment of probabilities to each branch. Except for the first branch, the probabilities are conditional on preceding branches. The use of conditional probabilities is one of the strengths of THERP, because is provides an analysis of task interactions.

583. The task analysis also includes the identification of performance shaping factors (PSFs). PSFs account for any factor that influences operator performance. The two basic classes of PSF are external and internal. External PSFs are sub-divided into situational characteristics (such as architectural characteristics, environment, shift system, supervision), task and equipment characteristics (such as perceptual, mental, memory, physical requirements, and human engineering factors), and job and task instructions (such as procedures, communications, and plant policies). PSFs include psychological stressors (such as high risks, vigilance, sensory deprivation, distractions), physiological stressors (such as fatigue, discomfort, pain, lack of exercise, and disruption of circadian rhythm), and organismic factors (such as experience, training, motivation, physical condition, and group identifications).

584. The most recent version of the technique (Swain and Guttman 1983) includes several "human performance models". These include the effects of psychological stress, vigilance, and the probability of diagnosing the system fault(s) within a certain time after the occurrence of an abnormal situation. Swain and Guttman (1975, 1983) note that although THERP was originally described as a human reliability model (in the sense of a set of relations and operating principles), it now seems preferable to restrict the use of the term to the human performance models incorporated in the overall technique.

History and Source

585. THERP was developed by Swain at Sandia National Laboratories in the early 1960s (Swain 1963). According to Swain and Guttman, the approach is an extension of studies conducted in the 1950s to estimate the influence of "first-order human failure terms" on the reliability of military systems. Swain developed his technique from a quantitative approach to the reduction of human error in industrial production, developed by L.W. Rook at Sandia.

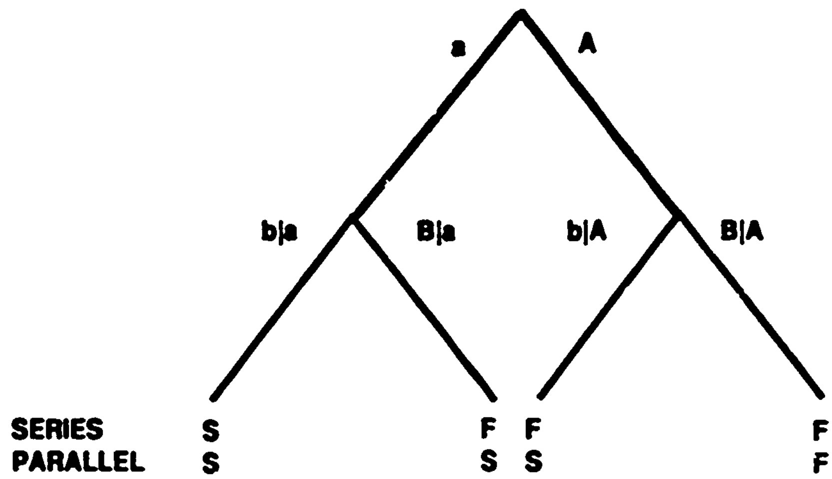


Fig. 4.11 Example of an HRA event tree, from Swain & Guttman, 1983

Let the first task be "A"; let the second task be "B"

a	=	probability of successful performance of task A
A	=	probability of unsuccessful performance of task "A"
b|a	=	probability of successful performance of task "B" given a
B|a	=	probability of unsuccessful performance of task B, given a
b|A	=	probability of successful performance of task "B" given A
B|A	=	probability of unsuccessful performance of task "B" given A

Probabilities of success S and failure F:

for "A" and "B" in series,

Pr[S] = a(b|a)
Pr[F] = 1 - a(b|a) = a(B|a) + A(b|A) + A(B|A)

for "A" and "B" in parallel
Pr[S] = 1 - A(B|A) = a(b|a) + a(B|a) + A(b|A)
Pr[F] = A(B|A)

586. THERP has been steadily developed and refined, particularly with regard to the range of tasks for which error probabilities are provided. The most recent version, and the most complete description, was published for the US Nuclear Regulatory Commission in 1983.

Product and Purpose

587. THERP produces two measures of performance:

(1) Task Reliability - an estimate of the probability that a task will be completed successfully;

(2) Recovery Factors - estimates of the probability of detecting and correcting incorrect task performance in time to avoid undesirable consequences on system performance.

588. According to Swain and Guttman, the purpose of the technique is to permit analysts to make quantitative or qualitative assessments of the probability of human errors that affect the availability or operational reliability of systems, and to permit the user to identify equipment designs, procedures, practices and other human factors problems which are likely to induce error. In addition to producing estimates of task reliability, THERP can be used to generate quantitative estimates of the interdependence of human activities, the effect of performance shaping factors such as training, stress, and fatigue, and the impact of equipment configuration and other system influences.

589. Although viewed by some as an hypothetical model, THERP was intended as a practical, applied technique, capable of providing systems designers and analysts with quantitative estimates of the effects of human errors on system performance.

When Used

590. Because it requires detailed information about operator tasks, the technique is best suited to well documented procedural tasks. Therefore, it is appropriate for the detailed design stage of systems development, and for the development and/or revision of operating procedures (sometimes referred to as "equipment procedure development").

Procedures for Use

591. The steps of the technique are similar to conventional equipment reliability analyses, except that they emphasise human operator activities(see Swain and Guttman, 1983; Sharit, 1988). The steps are:

(1) define the system failures of interest which may be influenced by human errors, including system goals and functions, and situational characteristics;

(2) list and analyse the related human operations, including the characteristics of the personnel, their jobs and tasks, performing a task analysis, and identifying error-likely situations (ELSs);

(3) estimate the relevant error probabilities, including the probability that the error will be undetected, or uncorrected;

(4) estimate the effects (consequences) of undetected errors on the system failure event analysis (this may require expert input from system reliability experts);

(5) develop recommended changes to the system, and recalculate the system failure probabilities.

Advantages

592. THERP was, and is, generally recognised as a significant advance over previous error analysis techniques, which merely assigned probabilities of error to discrete tasks. THERP permits the examination of the interdependence of tasks, and is compatible with engineering reliability techniques, thus permitting the combination of THERP analyses with other engineering analyses (Swain and Guttman, 1975).

Limitations

593. The most obvious limitation of the technique, like other reliability techniques, is that it is labour intensive. It requires a detailed operator task analysis if it is to be used effectively. For the system failure event analysis to be exhaustive, all possible operator errors must be identified as binary events. As Wagenaar has argued (1984), it is extremely difficult to anticipate all the errors that human operators might make. For this reason, some users concentrate on conducting a system fault tree analysis, and use THERP to study only those tasks associated with critical system faults. (This is, essentially, a more thorough version of the first step described in Procedure for Use, above).

594. An additional limitation is that the technique appears to require much judgment on the part of the user (designer), for the selection of appropriate error values, the degree of independence/interdependence among tasks, the selection of performance shaping factors, etc. Swain and Guttman (1975, 1983) address this point. They also admit to the problem of the lack of data on human error, and the difficulties of gathering data which are typically on the order of tenths of a percent.

595. Another criticism made of THERP concerns its use of simple models relating human performance to performance shaping factors (PSFs). The models are often so elementary and general that their validity can be questioned. The Handbook of Reliability Analysis (Swain and Guttman 1983) provides a broad collection of human error probabilities, but some estimates are based solely on expert judgment; others are extrapolations between data points. One example is the relationship between the probability of correct diagnosis of a plant fault and the amount of time following recognition of the abnormal situation.

Application Examples

596. THERP appears to have been widely used, although few examples have been published. Swain and Guttman (1983) and Bell and Swain (1983) report several cases of use and validation of the technique. In one case two analysts independently performed human reliability analyses (HRAs) for a system, using THERP. Few disagreements were found between the estimates, and they were resolved by a voting procedure. The final estimates were found to be within the 90% confidence interval of the actual production errors. Swain (1982) reports data from a validation study of errors in the installation of electronic components.

Technical Details

597. THERP is a manual process. Swain and Guttman (1983) provide human error probabilities for tasks, performance shaping factors, and details of the methods for analysis and quantification of human performance. The same reference also includes several specific models (some of them being similar to performance shaping factors). Such models include task dependence, response to displayed information, response to abnormal events, errors of commission in the use of manual controls, response to oral instructions and written materials, administrative control, level of stress, control room staffing, and recovery factors (the probability that an error will be corrected).

References

Bell, B.J. & Swain, A.D. (1983). A procedure for conducting a human reliability analysis for nuclear power plants. (Sandia National Laboratories, NUREG/CR-2254). Washington, DC: US Nuclear Regulatory Commission.

Miller, D.P. & Swain, A.D. (1987). Human error and human reliability. In G. Salvendy (Ed.) Handbook of human factors. (pp. 219-250). New York: Wiley & Sons.

Sharit, J. (1988). A critical review of approaches to human reliability analysis. Int. J. Industrial Ergonomics, 2, (111-130).

Swain, A.D. (1963). A method for performing a human factors reliability analysis. (Monograph SCR-685). Albuquerque, New Mexico: Sandia Corp.

Swain, A.D. (1982). A note on the accuracy of predictions using THERP, Human Factors Society Bulletin, 25, (4) 1-2.

Swain, A.D. & Guttman, H.E (1975). Human reliability analysis applied to nuclear power. Proceedings Annual Reliability and Maintainability Symposium. (pp. 110-115).

Swain, A.D. & Guttman, H.E. (1983). Handbook of human reliability analysis with emphasis on nuclear power plant applications. Final report. (SAND80-0200). Washington, D.C: U.S. Nuclear Regulatory Com ssion. (NUREG/CR - 1278F).

Swain, A.D. (1984). Qualifications of human error in LP/HC risk analysis. In: R.A. Waller & V.T. Covello (Eds.). Low-probability high-consequence risk analysis. New York: Plenum Press.

Wagenaar, W.A. (1984) Models of human operator failure. Soesterberg, The Netherlands: Institute for Perception TNO.

Future Needs

598. Swain and Guttman (1983) discuss potential developments of the technique, including revision of the current handbook, expansion of the current limited coverage of cognitive aspects of behaviour, and improvements on the accuracy of some of the estimates of

error. Swain (1984) argues that only the lack of good data has limited the usefulness of HRA methods such as THERP, and that (then) current efforts to remedy deficiencies in performance data should facilitate a better appreciation of the causes of human error. At the time of writing it is not known how many of those developments have been implemented.

4.6.2 Siegel-Wolf Models of Human Error

Summary Description

599. The human reliability models of Siegel and his colleagues can be seen as an outgrowth of earlier human performance modelling work (Siegel and Wolf 1961, 1969). The Siegel-Wolf model is a stochastic, digital simulation model of human performance which relates time to perform sub-tasks, time-stress and probability of successfully completing the task. It is reviewed in detail in sections 4.3.8 and 5.3.1 of this report. Siegel, Wolf and Lautman (1975) reported the development of six models which appear to be variants of the basic model (Siegel & Wolf 1961, Siegel, Wolf, & Williams 1976) which are applicable to human error analyses of system design. Those models appear to have been incorporated into the US Navy Human Reliability Prediction System User's Manual (Naval Sea Systems Command 1977). The models are:

(1) a hand calculation model for probability of correcting an equipment malfunction,

(2) a one-two man model for calculating malfunction correction time,

(3) a human and system reliability model for 4 to 20 men,

(4) an empirical model of equipment repair time (developed by Tracor Inc.),

(5) an "allocation model" which takes the outputs of the above models to maximise the reliability of man-machine systems using "standard dynamic programming techniques".

600. More recently, Siegel and his co-workers developed a computer-based Maintenance Personnel Performance Simulation (MAPPS), for estimating maintenance performance reliability in nuclear power plants (Siegel et al. 1984). The simulation calculates the probability of success of a sequence of maintenance sub-tasks, using a power function based on the difference between the sub-task difficulty and the maintainer's ability. The simulation includes ten classes of Performance Shaping Factors (PSFs), covering operator abilities, fatigue, and sub-task requirements, through to stress and organisational climate. Default values of PSFs are included. When the simulation is run, the probabilities of sub-task completion are weighted and summed to give an overall probability of success.

History and Source

601. The earliest attempts at producing a stochastic model of human performance (Siegel & Wolf 1961) included the interaction between time stress and task completion success. Therefore, the model is well suited to the analysis of tasks for which initial estimates of probability of task completion, and mean and standard deviation of task time can be produced. The basic structure of the Siegel-Wolf model is described in their 1961 paper. Descriptions of the five models developed for the US Navy are provided in the US Navy Manual (Naval Sea Systems Command 1977), available from the US Defense Technical Information Center. Details of MAPPS are provided in the three US NUREG reports.

Product and Purpose

602. The models in the US Navy Reliability Manual are directed at estimating the reliability of maintenance operations per se. The family of models permits the calculation of such factors as human reliability of correcting a specific malfunction, effect of time-stress on the reliability of maintenance technician performance, time to repair, utilisation of technician's time, effectiveness of one versus two technician teams, effect of allocation of maintenance technicians on overall system availability, etc.

603. The MAPPS model produces a calculation of the average probability of success of an operator (or crew) performing specific maintenance tasks. It also produces estimates of time to completion, areas of overload, idle time, and level of stress. The probability of success and overload information is used to examine the details of the task, or tasks, and the effects of changes in task details, including PSFs. Thus the model is intended to be used iteratively.

When Used

604. Because of the need for detailed task analysis data, the models appear best suited to the final stages of concept development, or to manning (Manpower, Personnel, Training, Effectiveness - MPTE) studies, or to the preliminary design stage. This applies to both the US Navy Human Reliability Prediction System, and MAPPS.

Procedures for Use

605. Each model requires different inputs, and one is a hand-calculation method. In each case the user must conduct a task analysis to determine what tasks are being performed, and then determine appropriate input factors, including those dealing with mental/physical activity, extent of instruction and supervision, use of reference materials, performance index ratings for individual technicians, probability of success of completing an individual task, and average task time.

606. A computer then simulates the performance of each subtask using the model algorithms and a Monte Carlo technique of simulation. The output of the simulation includes probability of success, time to completion, areas of operator overload, idle time, and level of stress. Changing parameter values and reiterating the simulation can demonstrate the effects of a particular parameter or subtask performance. Using this iterative method, the analyst can forecast the results of a potential design improvement prior to implementation. The input parameters are not treated independently, but interactively, to determine their collective effects on subtask performance. The following performance shaping factors are quantified by algorithms internal to the simulation and the input variables specified by the analyst:

(1) Maintainer's and work crew's abilities in terms of intellectual capacity and perceptual motor abilities.

(2) Fatigue effects, expressed as performance decrement due to number of hours of performance. Recovery in the form of rest is considered to lessen the fatigue level.

(3) Heat effects, considered as having a moderating effect on intellectual and perceptual-motor abilities.

(4) Subtask ability requirements by type of maintainer (e.g., maintenance mechanics, electricians) are identified including assembly, disassembly, and communication.

(5) Accessibility values for tasks such as removing and replacing components.

(6) Clothing impediment to perceptual-motor ability based on the interaction between accessibility and subtask difficulty.

(7) Quality of maintenance procedures.

(8) Stress effect, based on four stressors:
 (a) Time stress, the ratio of needed time to available time.
 (b) Communication stress, the percent of message comprehension as a function of ambient noise and message length.
 (c) Radiation stress, the stress as a linear function of radiation dosage beyond 800 mrems.
 (d) Ability difference stress, the maintainers' ability differences within the work crew.

(9) Aspiration level of the individual based on the ratio of successfully completed subtasks to the actual number of subtasks attempted.

(10) Organisational climate, policies, administrative structure, andvalues affect the detection of errors and their recovery.

Advantages

607. The Human Reliability Prediction System family of models permits a thorough analysis of the maintenance effort required to achieve a given level of operational readiness in large systems, such as naval vessels.

608. The MAPPS model permits the analysis of the effects of the ten performance shaping factors on maintenance task performance. One advantage of the model is that is treats the various parameters interactively, rather than independently, to determine their collective effects on sub-task performance

609. Both types of model are potentially useful for trade-off studies of maintenance tasks during systems design/development.

Limitations

610. The Human Reliability Prediction System family of models addresses only maintenance operations; it does not address other aspects of system operation, and would require a significant amount of effort to be used for other applications such as Operations Room (Combat Information Centre) system design. In such cases it might be more effective to

start with the basic Siegel-Wolf model and adapt it to the application.

611. Miller and Swain (1987) suggest that the "opacity" of the MAPPS simulation algorithm weakens its face validity, and could lead to a lack of confidence on the part of the analyst.

Application Examples

612. The models developed for the US Navy were based on data gathered from the analysis of existing systems, such as radar and sonar systems. The one-two man simulation model was validated using data from two shipboard radio systems (Naval Sea Systems Command 1977). No examples of applications are cited.

613. Applications of MAPPS are discussed by Siegel et al (1985).

Technical Details

614. The US Navy Reliability Manual provides some details of the programme variables and codes for the one-two and four-twenty man models.

References

Miller, D.P. & Swain, A.D. (1987). Human error and reliability. In: G. Salvendy (Ed.). Handbook of human factors. New York: Wiley & Sons.

Siegel, A.I. & Wolf, J.J. (1961). A technique for evaluating man-machine system designs. Human Factors 3, (1) 18-25.

Siegel, A.I., & Wolf, J.J. (1969). Man-machine simulation models: Performance and psychosocial interaction. New York: J. Wiley.

Siegel, A.I., Wolf, J.J., & Lautman, M.R. (1975). A family of models for measuring human reliability. Proceedings of the 1975 Annual Reliability and Maintainability Symposium. (pp. 110-115).

Siegel, A.I., Wolf, J.J. & Williams, A.R. (1976). A model for predicting integrated man-machine systems reliability. Wayne, PA: Applied Psychological Services Inc.

Siegel, A.I., Leahy, W.R., & Weisen, J.P. (1977). Applications of human performance reliability evaluation concepts and demonstration guidelines. Wayne, PA: Applied Psychological Services Inc.

Siegel, A.I., Bartter, W.D., Wolf, J.J., Knee, H.E., & Haas, P.M. (1984). Maintenance personnel performance simulation (MAPPS) model: Summary description. (Report NUREG/CR-3626, Vol 1). Washington, D.C.: US Nuclear Regulatory Commission.

Siegel, A.I., Bartter, W.D., Wolf, J.J., & Knee, H.E. (1984). Maintenance personnel performance simulation (MAPPS) model: Description of model content, structure, and sensitivity testing. (Report NUREG/CR-3626, Vol 2). Washington, D.C.: US Nuclear Regulatory Commission.

Siegel, A.I., Wolf, J.J., Bartter, W.D., Madden, E.G. & Kopstein, F.F. (1985). Maintenance performance simulation (MAPPS) model: Field evaluation/validation. (Report NUREG/CR-4101). Washington, D.C.: US Nuclear Regulatory Commission.

Naval Sea Systems Command (1977). Human reliability system user's manual. U.S. Department of the Navy. (AD 058 568).

Future Needs

615. The Human Reliability Prediction System models employ empirical data obtained during the early 1970s. Therefore, it seems necessary to examine the extent to which changes in the reliability of hardware and software, and changes in fault location and remove/ replace philosophies, require changes to the models or their supporting data.

This page has been left blank intentionally.

CHAPTER 5

MULTI-OPERATOR MODELS

5.1 INTRODUCTION

616. In the history of human operator model development, the authors view the multi-operator model development and applications as representing a transition period where interest was shifting from human operator models to human operator modelling technologies (reviewed in Chapter 8). Multi-operator models tend to be caught in a tug of war between the validity issues which dominate development of the elemental models of part-task or individual task performance, and utility issues behind modelling technologies such as SAINT. Validity issues focus on the extent to which the functional relationships within the model's structure represent reality, thereby assisting in the understanding of human behaviour. Utility issues deal mainly with how well the model helps the system designer conduct tradeoff analyses. The multi-operator models tend to still maintain considerable interest in the kinds of validity issues that characterise the more elemental models. However, the scope of multi-operator models has expanded far beyond our technical ability to provide a valid description of human performance for the variety of tasks and operators covered by the model. Thus, multi-operator models must also be judged in terms of their utility to the system designer. In this sense, they take on the characteristics of the modelling technology which is primarily a utility oriented approach.

617. It appears that in the struggle to develop multi-operator models that were both valid and useful, the emphasis shifted more in the direction of modelling capabilities such as SLAM and Micro SAINT. It was almost as if there was the recognition that one could not simultaneously balance valid individual/team performance predictions with the utility demands behind most requirements for individual/team models. In many instances, the individuals represented a crew such as a flight team which immediately translated to the need to specify overall system performance as related to individual/team behaviour. It is predicted that the modelling technologies will achieve the kind of widespread acceptance that has escaped most other models.

5.2 OVERVIEW AND RECOMMENDED REFERENCES

618. Multi-operator models appear to have been developed to satisfy two critical requirements in the system design arena. The first need is to estimate performance degradation experienced by team or crew members due to such conditions as stress, overload, or sustained continuous operations with their accompanying fatigue effects. Without a doubt, the Siegel-Wolf class of models reviewed in this chapter has been the most successful both in terms of breadth of applications as well as attempts to validate the models. Few other models can claim such an extensive validation effort.

619. The second need is the determination of the optimum crew size for a given weapon system. A number of crew performance models are reviewed in this chapter. They differ in terms of specific applications but follow the basic approach of taking a set of tasks and task time data and determining the best way to allocate the tasks to crews of varying sizes. Many of these models have been developed in the last few years. It is too early to tell whether there will continue to be a proliferation of such models or whether there will be consolidation around the most powerful, efficient, and easy to use models.

5.2.1 References

Meister, D. (1985). Behavioral analysis and measurement methods. New York: John Wiley & Sons.

Pew, R. W., Baron, S., Feehrer, C. E., & Miller, D. C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation (Report AFOSR-TR-77-0520). Bolt, Beranek and Newman, Inc., Cambridge, MA, (AD A038597).

Siegel, A. I., & Wolf, J. J. (June 1981). Digital behavioral simulation state-of-the-art and implications. (ARI Research Product 81-32). (AD128641)

5.3 MODEL SUMMARIES

5.3.1 Siegel-Wolf Model

Summary Description

620. The Siegel-Wolf model has been one of the most popular man-in-the loop simulations used by system designers. The model simulates maintenance or operator tasks simply by identifying personnel as operators or technicians, and the tasks as operator or maintenance tasks. It predicts task completion time and the probability of successful task completion. It also seeks to determine whether or not an average operator will successfully complete required tasks, how success probability changes for various performance shaping factors, and the operator proficiency required by the system. The model is both a design aid and an experimental tool. When the model is exercised with constant parameters, it is used as a design aid; when system parameters are varied on succeeding runs, the model is used in an experimental problem-solving mode. The basic assumption is that operator loading is the key element in effective man-machine system performance. A 1983 expansion of the model incorporates a probability of task success which depends on a comparison between the abilities required for successful subtask accomplishment and the maintainer's actual abilities. The inclusion of what is essentially a task difficulty factor significantly expands the model's capability of representing the mechanisms responsible for task accomplishment or failure.

History and Source

621. The Siegel-Wolf model was initially developed in 1960 at the Applied Psychological Services, Inc., Science Center by Arthur I. Siegel and J. Jay Wolf. It has been expanded over the years and still enjoys frequent use today. It is one of the most popular multi-operator models.

Product and Purpose

622. The purpose of the model is to serve as a tool for system designers during development, and to indicate where the system may over- or underload its operators. The model predicts task completion time and the probability of successful task completion. The model is both a design aid and an experimental tool. Model outputs include total number of runs (i.e., cycles through the task sequence), number and percent of successful runs, average time used over total runs, average time per run, average waiting time, average peak and final stress, the number of times a subtask was failed or ignored, the time spent in repeating failed subtasks, and the average time that the subtask was completed.

When Used

623. The Siegel-Wolf model can be utilised during the concept exploration through full scale development phases of the system acquisition process. The model is used mainly to provide equipment designers quantitative output for many workload related questions with particular emphasis given to the effects of stress on operator performance.

Procedures for Use

624. To exercise the model, 11 items of analytic input data are needed for each subtask and operator/technician. These identify: (1) decision subtasks; (2) non-essential subtasks which can be ignored in urgent conditions; (3) subtasks which must be completed by a second operator before it can be attempted by another operator; (4) time before which a subtask cannot be started; (5) the number of the subtask that must be performed next, assuming the current subtask is completed successfully; (6) the number of the subtask that must be performed next assuming the current subtask is failed; (7) average time in seconds required by the operator performing a subtask; (8) average standard deviation for the average operator; (9) the probability that a task will be performed successfully; (10) time required to perform all remaining essential subtasks; and (11) the time required to perform all remaining nonessential subtasks. The last two data input items are average execution times and assume no failures.

Advantages

625. The Siegel-Wolf modelling approach has wide applicability, has a broad recognition and acceptance among the scientific community, and has been strengthened in recent years by the inclusion of a task difficulty factor.

Limitations

626. The model's complexity suggests that its use is limited to large scale system development applications. As with most network type models, there is a requirement to collect descriptive data of the system being modelled. Therefore there is a fairly substantial front end investment in terms of time and labour. The models have been difficult to use by other than the developer and difficult to create using published documentation. The fact that the model is mainly sequential requires operators to complete tasks in the assigned sequence.

Application Examples

627. MAIN - Simulation of Human Performance in Electronic Imagery Systems. The programme simulates ground-based processing of targets as viewed by photo-reconnaissance operators. Data are obtained by airborne side-looking radars. Programme outputs include operator time required for task performance and adequacy of operator's task performance for five different modules: scan/detect, classification, alternate classification, decision, and communication. The programme has not been validated. Additional information regarding MAIN is contained in Siegel, Wolf, and Williams (1978).

628. NETMAN - Simulation of Message Processing in Military Exercise Control and Evaluation Systems. The programme simulates information processing by Army personnel in field exercises and tests the effectiveness of a system composed of up to three networks. It calculates information loss and an effectiveness index composed of thoroughness, completeness, and responsiveness measures. (See Siegel, Leahy, and Wolf, 1977). NETMAN has been partially validated. (See Siegel, Leahy, Wolf, and Ryan, 1979).

629. Intermediate Size Crew Model. The model simulates man-machine systems operated by crews of 4 to 20 members. Scheduled event data include such inputs as: duration, essentiality, crew size, energy consumption, etc. Input data examples for

equipment and repair events are: failure rate, repair time, mental load. Examples of input data for emergency events are: probability of occurrence, recovery time, and expenditure rate of consumables. Personnel input data examples are: percentage of fully qualified crew members, cross training probability, average stress threshold, sleep rate per day, and daily caloric intake.

630. In addition to providing output data for the various input parameters, the programme provides crew pace data showing the effect of speed of personnel on the number of successful events completed. Data can show the effect of stress on successful and failed events, the effect of crew qualification on mission results, the effects of a low supply of expendables (e.g., water and fuel) on mission results, the length of the workday on the hours worked, and the effect of varying crew size on system performance.

631. An operational Viet Nam river patrol mission was simulated and validated by comparing the simulation results with those obtained from interviews conducted with Coast Guard officers. (See Siegel, Wolf, and Cosentino, 1971).

632. <u>One-and Two-Man/Machine Model</u>. The model determines if an operator can successfully complete the tasks within a given time limit, the success probability for slower and faster operators and shorter and longer time periods, the effect of stress on performance, and the distribution of failures as a function of operator stress tolerance and operator speed.

633. Input data include: type of subtask, subtask essentially, idle time between tasks, sequence of subtask performance, subtask execution time, subtask probability of success, time remaining to perform subtasks, goal aspiration, stress condition of operator, team cohesiveness.

634. The model generates the following data for each operator: stress threshold, time used, time available, time overruns, stress, cohesiveness and goal aspiration at the end of the iteration. The model produces graphic plots such as the probability of success as a function of speed, time available, and stress threshold.

635. Validation studies were accomplished for both the one-man and two-man simulation for several different types of mission (carrier landing, missile launching, in-flight refuelling). All results for the model's prediction ability were favourable. (See Siegel and Wolf, 1969).

636. <u>Group Simulation of Man-Machine Systems</u>. The model predicts qualities of larger (up to several dozen men) man-machine systems such as: systems efficiency, crew morale and cohesiveness, time devoted to equipment repairs, manpower time shortage as a function of crew size, proficiency of crew members, and manhour loadings and overtime.

637. Input data include: average time to complete action units, communications between stations to complete an action unit, importance of action unit, type of action unit (normal, training, difficult), equipment failure rate, average repair time, number of personnel required to operate equipment, probability of a crew member having one alternate and two alternate specialties, number of personnel types allotted to line and staff, crew size increments, morale threshold, working hours per day, probability of an emergency, proficiency for the average crew member, and crew member pay level.

638. Minimum crew size is determined from workload data; execution time from group member proficiency, morale, overtime load, number of men required, and average time. The model determines crew morale and cohesiveness, work groups, communications efficiency, proficiency of performance, action unit execution time, psychological efficiency, environmental efficiency, total efficiency, and performance adequacy.

639. Content and construct validity were tested through the simulation of typical missions and the model's results were compared with actual mission data. Favourable results were found. (See Siegel and Wolf, 1969.)

Technical Details

640. The following are the basic parameters of the model:

641. The parameter T_j, the mission time limit, specifies the total time allotted to each operator for performance of the task.

642. The parameter F_j accounts for variance among individuals operating the system. This parameter enables the model to simulate an operator who usually performs faster or slower than the average operator. The effects of faster, or more highly motivated operators (F_j>1), and slower operators (F_j<1) are examined by performing several computer runs with different F_j values.

643. A third parameter which is central to the model is the stress threshold M_j, operationally defined as the ratio of how much is left to do to the amount of time available in which to do it.

644. The critical importance of stress is indicated by its relationship to probability of successful performance of the subtask P_{ij}. Thus the probability of success increases linearly with stress until it becomes unity at the stress threshold, after which the probability decreases linearly until, when stress has a value equal to $M_j + 1$, it levels off at a value which is decreased from P_{ij} by an amount equal to $P_{ij}/2$.

645. Similarly, execution time for the subtask varies as a function of stress. If the average operator requires T_i seconds to perform subtask i when stress is unity, T_i decreases with increasing stress until M_j is reached after which T_i increases linearly with increasing stress.

646. The model was written in FORTRAN and has enjoyed a variety of applications on many different systems including the IBM 30-33. For additional information on the model, contact Applied Technical Associates, 345 Robinson Dr., Broomall, PA 19008.

References

Siegel, A. I., Leahy, W. R., & Wolf, J. J. (1977, October). A computer model for simulation of message processing in military exercise control evaluation systems (ARI Technical Report TR-77-A22) (AD-A045832)

Siegel, A. I., Leahy, W. R., Wolf, J. J., & Ryan, G. R. (1979). Application of computer simulation techniques in military exercise control system development: 1. NETMAN model sensitivity test and validation (ARI Technical Report 407). (AD A081993)

Siegel, A. I. & Wolf, J. J. (1961). A technique for evaluating man-machine system design. Human Factors, 3 (1), 18-28.

Siegel, A. I. & Wolf, J. J. (1962). A model for digital simulation of two-operator man-machine systems. Ergonomics, 5, 557-572.

Siegel, A. I. & Wolf, J. J. (1969). Man-machine simulation models: Psychosocial and performance interaction. New York: Wiley.

Siegel, A. I., Wolf, J. J., & Cosentino, J. (1971). Digital simulation of the performance of intermediate size crews: Application and validation of a model for crew simulation. Wayne, PA: Applied Psychological Services.

Siegel, A. I., Wolf, J. J., & Williams, A. R. (1978). Computer simulation of human performance in electronic processed imagery systems. Wayne, PA: Applied Psychological Services.

Future Needs

647. More routine use of the model should surface future needs.

5.3.2 Models of Operator Performance in Air Defense Systems (MOPADS)

Summary Description

648. MOPADS is a set of modelling tools to represent the performance of human beings in complex man-machine air-defence systems. Three types of operator activities are represented in MOPADS. The first is skill-based behaviour. Skill-based behaviours are simple control actions such as a button press. The second type of behaviour represented in MOPADS is rule-based behaviour. This type of behaviour is typified by the performance of check lists. The third type behaviour is knowledge-based behaviour. This type of behaviour is strategic in nature. MOPADS operators evaluate the potential impact of available operator tasks on their goals when selecting the next task to perform. They engage in goal seeking behaviour in which they either maximise the expected reduction in the average goal priority or maximise the expected reduction per unit time of the average goal priority. The approach taken in MOPADS has been to consult subject matter experts (in addition to Army documentation) to determine a sufficient set of operator goals.

649. Human factors modules developed for MOPADS are stand-alone software modules. Alternate human factors representations can be readily tested within the MOPADS system.

650. The simulation methodology selected for MOPADS is discrete event simulation. The SAINT simulation language has been selected as the host language for the MOPADS operator models. SAINT provides a formal capability to introduce human factors considerations by using "moderator functions" which modify the nominal time to perform tasks.

History and Source

651. MOPADS was developed by the U.S. Army Research Institute Field Unit - Fort Bliss, Texas, under contract to Pritsker & Associates, Inc. with a sub-contract to Calspan Corp. The period of performance was from 1980-1983.

Product and Purpose

652. Products include (1) MOPADS Final Report with twenty-seven supporting appendixes describing various aspects of the model and its use, and (2) MOPADS software. A separate document has been prepared which contains user instructions on how to use the MOPADS User Interface to set up and perform MOPADS simulation. It also contains information on analysing the outputs from simulation.

When Used

653. MOPADS was intended to be used to model operator performance and relate it to system performance for the AN/TSQ-73 and IHAWK systems. The moderator functions permit the investigation of the degrading effects of sustained operations by representing the influence of operator characteristics (i.e., skill level) and environmental characteristics (i.e., trackload).

Procedures for Use

654. The procedures for use are quite complex and not easily summarised. A detailed description can be found in the user's guide in the previously referenced 27 appendices which accompany the final report.

Advantages

655. The major advantages of MOPADS are (1) the ability to network systems such as the AN/TSQ-73 and IHAWK and examine their interaction, and (2) the ability to portray certain human performance decay functions resulting from sustained, continuous operations.

Limitations

656. The major limitation is the size/complexity. It shares many of the same limitations of SAINT in this respect. While it purports to be user friendly, it probably would require a software support centre to maintain it. The rapidly developing micro-computer models are discouraging users from investing the resources necessary to exploit the many attractive features built into MOPADS. The models are primarily for air defence applications.

Application Examples

657. The original intent was to validate MOPADS using actual AN/TSQ-73 and IHAWK systems, but this goal was not achieved.

Technical Details

658. MOPADS has a skills taxonomy consisting of 19 skills. In addition there are 64 operator state variables, 9 environmental variables, and 11 task related variables. MOPADS supports the following distribution functions:

(1) Constant

(2) Normal

(3) Uniform

(4) Erlang-1 (Exponential)

(5) Lognormal

(6) Beta

(7) Gamma

659. The MOPADS programme is written in FORTRAN. A separate document has been prepared which describes the FORTRAN style and documentation requirements. All new FORTRAN code written for the MOPADS project has been written following the standards specified in this document. The programme was written for use on a VAX 11/730.

Questions concerning acquisition of the software or documentation should be sent to ARI Field Unit, ATTN: R&D Coordinator, P.O. Box 6057, Fort Bliss, Texas 79906-0057.

References

Polito, J. and Laughery, K. R. (1986). MOPADS final report (ARI Research Product 84-144 and 84-145). (ADA162600 and ADA162880)

Future Needs

660. Simplification is needed such as converting the model for use in a micro-computer. Further validation is also required.

5.3.3 Performance Effectiveness of Combat Troops (PERFECT)

Summary Description

661. PERFormance Effectiveness of Combat Troops (PERFECT) is a computer simulation model of the degradation of combat effectiveness and stress build-up of combat troops during periods of continuous operations. The model is basically a force-on-force, theatre level type of simulation. It simulates groups/teams/units performing combat activities. The model represents nine combat unit types (i.e., 155mm Howitzer section, tank crew, fire support team) and the 16 duty positions (i.e., Howitzer section chief, 155mm gunner, tank commander, tank loader) of which these units are structured. The three combat activities simulated are: (1) repelling an attack from a battle position (fire mission); (2) creating and defending a strongpoint (emplacement); and (3) disengaging and occupying a new battle position (roadmarch).

662. PERFECT permits analysis of anticipated performance effectiveness when variables such as continuous time in battle, light level, enemy/ friendly numerical ratio, enemy/friendly terrain advantage, amount of platooning, and amount of sleep permitted are varied alone or in combination. The model is based on a series of manipulations of a four dimensional matrix of values of combat troop's effectiveness characterising crewman duty position, crewman tasks, combat day, and combat mission. These crewman matrices are aggregated into unit daily effectiveness values as a function of the simulated set of circumstances occurring throughout the combat day. Crewman task effectiveness values are also aggregated to determine combat effectiveness levels for 5 factor analytically derived factors: (1) command and control; (2) combat activity; (3) coordination and information processing; (4) preservation of forces and regrouping; and (5) orientation to friendly and enemy troops. Stress level is determined as a function of light level conditions, terrain advantage, squad proficiency levels, enemy/friendly personnel strength ratio, and enemy/ friendly material strength ratio.

663. The model is designed for interactive operation at a terminal by a user with no or minimum sophistication in computer science or computer use.

History and Source

664. The PERFECT model was a product of the U.S. Army Research Institute, Alexandria, Virginia, research on Human Performance in Continuous Operations. Earlier work was published in three volumes. Research was conducted as part of the Army project, Man-Machine Interface in Integrated Battlefield Control System, FY78 Work Programme and research was supported by Fort Leavenworth, Kansas, which was the U.S. Army Training and Doctrine Command (TRADOC) sponsor.

Product and Purpose

665. The PERFECT model was developed to aid in the understanding of human performance during night and continuous operations. It was designed as a means of assessing the cumulative effect of several stress-producing variables on human performance during continuous combat. The model allows insight on the potential interaction between variables which affect combat performance. Model outputs consist of three combat day summary tables estimating (1) combat unit effectiveness values, (2) combat activity

effectiveness values, and (3) combat unit maximum stress level values.

When Used

666. As presently configured, the model has utility for tactical planners and training specialists. The model can aid in the realisation of training needs, doctrine needs, equipment needs, and tactical requirements The model may also be used in conjunction with other war-gaming models. The model can be used as an "exercise" in planned training programmes, as a means of demonstrating the importance of human performance to mission success.

Procedures for Use

667. The user/analyst must prepare and input the following data:

(1) select the number of men in each operational unit to be simulated

(2) any desired changes in the basic effectiveness tables

(3) mission sequence data

(4) unit proficiency factor for each unit

(5) enemy/friendly ratios for material strength, personnel strength, and terrain advantage

(6) light level profile for each day

(7) unit replacement data

668. For more specific detail and step-by-step terminal user's instructions, see Siegel, Wolf, Schorn, and Ozkaptan (1981).

Advantages

669. The major advantages of the PERFECT model are (1) it handles operations up to five days in duration, (2) it considers a maximum of 16 duty positions and nine combat unit types, and (3) it simulates three platoon actions for which data are currently available on the basis of five summary factors. The model is designed so that these limits can be expanded later, if desired, to simulate different scenarios and additional duty positions, combat unit types, factors, and platoon actions. Also, the model is designed for interactive operation at a terminal by users with little or no modelling experience.

Limitations

670. The most notable limitation of PERFECT relates to the fact that the model simulates unit level considerations: units degrade as a group, get stressed and relieved as a group, sleep as a group, etc. Also, when a unit is replaced, it is by a "fresh" unit, i.e., a non-degraded unit of personnel. The consequence of unit level manipulations is that PERFECT does not assist in the design or in the evaluation of the design of either the hardware or the crew. For example, function allocations and/or tasks assignments cannot be manipulated.

Application Examples

671. The model is intended to model military continuous operations. Applications have been limited. The technical report does contain sample specifications of baseline input, and predicted outcomes from various hypothetical missions.

Technical Details

672. Specific detail about the functions, logic, processing, and calculations which the programme performs are provided in Appendix B of the referenced technical report. The model is hosted on the UNIVAC 1100 system and is compiled on the FORTRAN (ASCII) compiler.

References

Pfeiffer, M. G., Siegel, A. I., Taylor, S. E., & Shuler, L. (1979). Background data for the human performance in continuous operations guidelines. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Siegel, A. I., Pfeiffer, M. G., Kopstein, F. F., Wilson, L., & Ozkaptan, H. (1979). Human performance in continuous operations: Vol. I. Human performance guidelines. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A086131).

Siegel, A. I., Wolf, J. J., Schorn, A. M., & Ozkaptan, H. (1981). Human performance in continuous operations: Description of a simulation model and users manual for evaluation of performance degradation (ARI Technical Report 505). (AD A101950).

Future Needs

673. The model is designed so that limitations can be expanded later, if desired, to simulate different scenarios and additional duty positions, combat unit types, factors, and platoon actions.

5.3.4 Simulation for Workload Assessment and Modelling (SIMWAM)

Summary Description

674. Simulation for Workload Assessment and Modelling (SIMWAM) is a microcomputer-based task network modelling technique used to assess operator workload and performance effectiveness in man-machine systems under alternative design and manning configurations. The tool performs simulation of the functioning of a man-machine system represented as a network of tasks. The network model is composed of tasks, resources to perform tasks (personnel, hardware, software), and relationships between the tasks. SIMWAM consists of a set of five related programmes which permit the analyst to create, maintain, and modify a data base of task requirements; to execute the task network; and to output simulation results. The software includes the capabilities to specify task priorities, to interrupt and restart tasks, to branch deterministically or probabilistically, and to interface user-written subroutines with the five SIMWAM programmes.

675. Three versions of SIMWAM are available: Automatic, semi-automatic, and manual (or interactive). The three differ primarily with regard to the locus of control of task sequencing. During execution of the automatic version of SIMWAM, processing is entirely controlled by the software. In the semi-automatic mode, the user manually controls only probabilistic branching. When probabilistic branching situations are encountered, the programme pauses and waits for the user to select the next task from the appropriate task subset. With the interactive version, the user manually controls all task sequencing. At points when task flow decisions are to be made, interactive SIMWAM presents a menu allowing the user to review current task and/or operator status, to perform operator assignments, to start tasks, to select task performance durations, and/or to exit from the programme.

History and Source

676. The technique was developed by Carlow Associates, Incorporated for the U.S. Navy Sea Systems Command in order to support the analysis of workloads and information flow in Naval carrier air operations. The impetus for development of SIMWAM was the need to analyse the potential impacts of manipulating manning levels and/or function allocations on system effectiveness and operator workload during the performance of aircraft detection and tracking tasks within the Combat Information Center of surface ships.

677. The SIMWAM software is maintained and distributed by Carlow Associates, Inc., 8315 Lee Highway, Suite 410, Fairfax, VA 22031-2269. Copies of the software can also be obtained through the U.S. Navy Sea Systems Command, SEA61R2, Washington, D.C.

Product and Purpose

678. The primary purpose of SIMWAM is to provide operator utilisation data in order to assess, redistribute, and reduce crewman workload levels. The tool enables the analyst to evaluate system design, or system modification, concepts involving personnel reduction, cross training, task modification, automation, function allocation, and different operational conditions or strategies. It provides comparative simulation outputs which can be used to

determine the efficacy of the alternative concepts. Outputs from the execution of the model can be obtained on:

(1) Sequence of task events showing task times (start, end, and duration), task interruptions, task completions, and operators assigned

(2) Summary of task completions and operator time on each task

(3) Matrix of the time spent by each operator on each task

(4) Summary of busy and idle times for each operator

When Used

679. SIMWAM is intended to be used to help assess and reduce workload and manning levels. The interactive version of the programme is particularly applicable to multi-operator systems in which the assignment of individual tasks to operators can be varied in real time to alleviate excess workload. It can also be used to evaluate various function allocation (man-machine-software) schemes. SIMWAM can be used during any phase of a system's life.

Procedures for Use

680. The user must create a data base of task information, execute the task network, evaluate the simulation results, then modify the task data base in order to evaluate alternate system concepts or configurations. The task information consists of:

(1) Task number and name

(2) Task analytic information concerning predecessor and successor tasks

(3) Task performance data (i.e., minimum, medium, and maximum times)

(4) Specification of the operator(s) qualified to perform each task and the number required

(5) Task priorities that control operator assignment to tasks

(6) Task interruption criteria in case of operator assignment conflicts

Advantages

681. The major advantage of SIMWAM is that it enables one to simulate sizable system models on a microcomputer. Also, the three execution modes (automatic, semi-automatic, and manual) provide a high degree of flexibility in reallocating tasks based on system demand and on operating conditions.

Limitations

682. A minor limitation of SIMWAM is that the software is currently only available for the Tandy Radio Shack TRS-80 and the Apple Macintosh microcomputers. Also, the triangular probability distribution is the only random sampling function contained in the programme.

Application Examples

683. The SIMWAM programme was used to develop and exercise models of activities and workloads in operations aboard U.S. Naval surface ships. Specific applications include the assessment of workloads of:

(1) 35 operators conducting flight operations within a carrier Aircraft Management Center

(2) 9 operators conducting activities in the surface/subsurface area of a shipboard Combat Information Center

(3) 10 operators performing detection and tracking operations within the Combat Information Center

Technical Details

684. The SIMWAM software is written in Microsoft BASIC and consists of approximately 3,000 lines of code. The programme is currently available on floppy diskettes for the Tandy Radio Shack TRS-80 (5.25" disk) and the Apple Macintosh (3.5" disk) personal computers. It has been validated as a workload assessment tool on Naval Shipboard Systems.

References

Kirkpatrick, M., Malone, T. B., & Kopp, W. H. (1984). Development of a model of shipboard detection and tracking operations using SIMWAM. (Final Report). Carlow Associates, Inc.

Kirkpatrick, M., Malone, T. B., & Kopp, W. H. (1985). Development of a model of shipboard surface/subsurface operations using SIMWAM. (Final Report). Carlow Associates, Inc.

Future Needs

685. There is a need for an IBM PC compatible version of software. Interest exists regarding this improvement and a potential work effort is under consideration by Carlow Associates. A further valuable modification to the programme would entail the incorporation of human performance degradation routines, especially in the area of sustained, continuous, combat operations.

5.3.5 METACREW

Summary Description

686. METACREW is a fast-time computer simulation of the Joint Surveillance Target Attack Radar (JSTARS) Ground Station Module (GSM). The GSM crew monitors and tracks ground targets such as supply, artillery, and tank units. The model simulates the crew's target processing and decision making tasks and the normal sequential flow of these tasks. Components of the model include: (1) the operator behavioral model, (2) the personnel model, (3) the battlefield scenario, (4) the commander's guidance, and (5) the output performance data file.

687. The operator model, which is a network representation of the operator's tasks, consists of system tasks (keypress actions), decisions, and processes. Processes are sequences of individual tasks and the associated operator decisions which result in the performance of tactical missions. This model of the operator is driven by the battlefield scenario which includes targets moving on the battlefield and special requests for information or for fire support. Scenario events are selected for processing by the operator model based on the commander's taskings and priorities regarding targets and special requests. This guidance plays a critical role in terms of mission emphasis and crew workload. Workload and system performance are also driven by characteristics of the GSM crew. The personnel model enables the analyst to describe and manipulate the crew in terms of (1) the number of operators, (2) operator task assignments, and (3) operator skill or performance level. Skill level is represented in terms of performance time distributions for the various system tasks. Data descriptive of operator and system performance during the simulation exercise are stored in the output file.

History and Source

688. METACREW was developed at Honeywell Systems and Research Center in 1985 for the U.S. Army Electronics Research and Development Command (ERADCOM). The simulation model was developed out of the need for a time-efficient tool that could examine the relationship between human variables and overall system performance under a wide range of battlefield conditions. The user's manual for METACREW and documentation for the SIMSCRIPT (Simulation Scriptor) programme are contained in Gilles (1982). Additional information regarding the software model may be obtained from Honeywell Systems and Research Center, 3660 Technology Drive, Minneapolis, Minnesota 55418 or from U.S. Army ERADCOM, Fort Monmouth, New Jersey.

Product and Purpose

689. METACREW provides a means of rapidly evaluating the impact of changes in the human operator's role in the JSTARS GSM. Products include a description of operator performance in terms of work histories, performance statistics, and mission summaries. An event history output contains a complete listing of each event processed by the system operators. The activity summary provides a statistical summary of each operator's performance time on each task and the target status summary contains the time required to process a target.

When Used

690. METACREW is specific to the JSTARS GSM, and therefore, it is only applicable to that system. However, it could fairly easily be modified to represent other data processing/sensor systems. It can be used to assess issues involving man-machine interface design, crew size, crew configuration, crew skill level requirements, performance style differences, and doctrine and tactics.

Procedures for Use

691. In order to use the simulation model the user must prepare and input:

(1) the crew size

(2) job/task assignments

(3) crewmen skill level or proficiency information in the form of performance time distributions

(4) the command rules

(5) characteristics of the scenario environment including targets, special requests, interruptions, and system failures

Advantages

692. The model is based on extensive data from live exercises and from man-in-the-loop runs of the JSTARS Ground Station simulator. Therefore, METACREW contains an accurate and detailed representation of operator tasks, decision logic, operator skill level characteristics, etc.

Limitations

693. The major drawback of METACREW is the fact that it is a specific model of the JSTARS GSM only and the empirical data base incorporated into the model is not generalizable to other systems due to software/hardware response characteristics. Also, the model does not accommodate erroneous or degraded performance.

Application Examples

694. METACREW has been used to evaluate different crew size configurations, to determine the tactical utility of the GSM as a function of tasking, and to assess alternative multiperson/multiple van networking configurations.

695. The METACREW simulation has been validated against the performance of experienced JSTARS data handling teams during military exercises (Plocher, Gilles, & Tamanaha, 1985). In these validation trials, the simulation was shown to account for 76-96% of the variance in the performance of the real operators. Further, the simulation was shown to respond to workload challenges in a manner similar to actual operators.

696. Since its validation, METACREW has been used to explore a variety of system development issues on the JSTARS programme. Its initial application was to identify crew configurations that avoided bottlenecks in the flow simulation. It has also been used to estimate the impact of new missions on crew performance.

Technical Details

697. The model is coded in SIMSCRIPT for the Digital Equipment Corporation VAX 11/780 computer.

References

Plocher, T. A., Gilles, P. L., & Tamanaha, R. J. (1985). METACREW: A simulation of GSM crew operations. (Preliminary Report). Honeywell Systems and Research Center.

Gilles, P. L. (1982). User's manual for METACREW. U.S. Army ERADCOM, Fort Monmouth, NJ.

Future Needs

698. The model would benefit from the inclusion of the capabilities to simulate erroneous and/or degraded performance. A version of the model generalizable to data processing/ sensor systems would also be of value.

5.3.6 Crew Performance Model (CPM)

Summary Description

699. The Crew Performance Model (CPM) is a computer-based, crew simulation model and performance evaluation tool. It is a Monte Carlo procedure for estimating the relative effects on crew performance of changes in crew size and task assignments. The concept underlying CPM involves task reallocation as a means of assessing alternate crew structures. The model consists of three components: A task library containing the relevant crew tasks, the task interdependencies (concurrences and dependencies) and the associated task performance times; an input programme for entering the crew size, crewman task assignments and task orders; and the main programme for integrating the input information (from the input programme and the task library) and calculating summary performance measures.

History and Source

700. CPM was developed by personnel at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) Fort Sill Field Unit in 1979. The initial phase of CPM development aimed to simulate crew performance under "optimal" conditions, i.e., no forms of human performance degradation were to be simulated. The objective of CPM was to provide a method for simulating and estimating the performance of alternate crew structures under conditions of degradation resulting from (1) continuous operations, (2) fatigue, (3) sleep degradation, (4) stress, (5) NBC environments and (6) night operations. The goal of incorporating degradation routines into CPM was never achieved. A complete listing of the CPM software is contained in Appendix B of Schwalm, et al. (1981).

Product and Purpose

701. The overall objective of CPM is to enable decision makers and research personnel to study, in a timely and cost-effective manner, the effects of varying crew size and task assignments without expending the time and dollar resources necessary to observe alternate crews actually performing system functions. The model computes and outputs: (1) the total time required for the entire crew to complete its mission or job, (2) a distribution of these mission completion times across iterations, (3) identification of the "critical man", or busiest crewman, on each iteration, with that status cumulated across iterations for each crewman, and (4) the percentage of idle time per iteration per crewman. These performance measures can be used for evaluating the speed and relative efficiency of crews varying in size or structure.

When Used

702. The model can be used to evaluate the effects of varying crew size and crew member task assignments on the performance of crew-served systems. It is also applicable to the evaluation of the effects of equipment changes to an existing or proposed hardware system. The tool can be used iteratively throughout the system development process.

Procedures for Use

703. The analyst must obtain task analysis and task performance data developed from time and motion analyses, from system documentation, or from expert estimates, etc. A system task library is created which contains the following information for each task characterising the man/machine system:

(1) Task number and verbal descriptor

(2) Distribution of task performance times consisting of a minimum, mean, and maximum time

(3) Hold, or dependent, task number

(4) Concurrent task number

704. The task library operates as an independent portion of the model; tasks, or task information, may be added or deleted as the need arises. Once a system task library is constructed the analyst/user specifies the input programme's data file. In order to model a particular crew structure, the user only needs to specify the number of crew members, and then to assign tasks, by task number, to individual crew members in the order in which the tasks are to be performed. The input programme essentially directs the operation of the model in simulating the performance of alternate crew structures based on the task information contained in the task library. Finally, the model is executed, and the main programme calculates and outputs performance information based on the task library and crew structure specifications.

Advantages

705. CPM is conceptually simple and straightforward. The input programme/ task library interface alleviates the need to modify the simulation model software when simulating different systems or crew structures, and the output is diagnostically useful. In addition, the tool is easy to use, cost effective, and provides for rapid output.

Limitations

706. Two major drawbacks to CPM are that (l) the capabilities to assess/estimate performance degradation due to continuous operations, etc., were never incorporated into the model and (2) the model cannot simulate or represent branch tasks. This latter limitation prevents an analyst from simulating decision behaviour, erroneous performance, and probability/criticality of task performance. In addition, tasks can be represented only as unidimensional, time varying entities, and such differences as the workload/fatiguing or cognitive/physical characteristics of tasks were not considered. Also, the triangular probability distribution is the only random sampling function contained in the programme.

Application Examples

707. The model has been applied to M109A1 Howitzer sections to determine the optimal crew size necessary to support operations while assuring an equatable distribution of workload among crew members. CPM has also been used to analyse the performance

capabilities of alternate Division-86 155mm Howitzer battery organisations.

Technical Details

708. CPM is written in FORTRAN and the main programme consists of approximately 550 lines of code.

References

Crumley, L. M. (1979). Tasks and possible task assignments for a ten-man crew, divided into two five-man units emplacing, firing, and march ordering an M109A1 howitzer section (ARI Fort Sill Field Unit Working Paper 79-4).

Coke, J. S., Crumley, L. M., & Schwalm, R. S. (1981). Emplacing, firing, and march ordering an M109A1 howitzer: Tasks and task times (preliminary) (ARI Research Report 1312). (AD A109706)

Schwalm, R. C., Crumley, L. M., Coke, J. S., & Sachs, S. A. (1981). A description of the ARI crew performance model (ARI Research Report 1324). (AD A113793)

Schwalm, R. C. & Coke, J. S. (1981). User's guide for the ARI crew performance model (ARI Fort Sill Field Unit Working Paper 82-1)

Crumley, L. M., Schwalm, R. S., & Coke, J. S. (1982). An evaluation of the effects of various task assignment alternatives on M109A1 howitzer crew performance (ARI Research Report 1337).

Robinson, R. E. & Hannon, C. W. (1982). An analysis of the capability of alternate Division-86 155mm howitzer battery organizations. (Final Report). Science Applications, Inc.

Future Needs

709. The model can be further developed to evaluate performance decrements (or increments) as a result of factors such as continuous operations, extreme temperatures, NBC conditions, crew turbulence, and training. Also, random sampling functions in addition to the triangular probability distribution could be incorporated into the software.

CHAPTER 6

BIOMECHANICS AND WORK SPACE DESIGN

6.1 INTRODUCTION: BIOMECHANICAL MODELS

710. Increasing importance has been attached to the health and safety aspects of manual materials handling tasks in both industrial and military occupations. It has been estimated, for example, that 25 to 30% of all overexertion injuries are caused by handling materials manually (Mital 1983). Despite increased automation, a high proportion of military operations and maintenance tasks exploit human adaptability in the manipulation or movement of large or heavy items. For example, a recent survey of 100 Canadian Armed Forces (CF) trades revealed that the most physically demanding tasks in almost every trade involved lifting. In the CF, 20% of all reported work-related injuries which resulted in absence from duty were back-related (Warrington-Kearsley, 1986).

711. The biomechanical models reviewed here attempt to predict such injuries through the application of the laws of mechanics to the description and analysis of human movement, strength, and lifting ability. Biomechanics has a long and diverse history. Applications have included bone strength, the study of individual joints, groups of joints, body movement, and lifting and carrying (Kroemer, Snook, Meadows, & Deutsch, 1988). Only models of lifting and carrying have been included in this review. Biomechanical modelling applications have also included the description of human response to vibration and shock. The latter kind of model has not been included in this review because, in general, they describe human tolerance rather than human performance. Note, however, that Coerman, Magid and Lange (1962), and others, have shown that such models can predict interference with the performance of some tasks.

6.2 OVERVIEW AND RECOMMENDED REFERENCES: BIOMECHANICAL MODELS

712. Biomechanical models of the human body must describe a complex, non-linear system. As Kroemer, Snook, Meadows, and Deutsch (1988) have noted, the study of even simple physical tasks can involve many more unknown muscle groups than there are independent equations to solve them. Possibly for that reason, most biomechanical models are limited to one or two joints. Of 99 models reviewed and tabulated by King and Marras (1988), only five were classified as "whole body". Nevertheless, a considerable amount of effort has been put into modeling human materials handling capabilities (Chaffin, 1985; Drury, 1976). Although such models appear to have been used most frequently in the investigation of existing tasks to alleviate problems, they are being used increasingly for operator task and work place design.

713. The analysis of a manual handling task must take into account three classes of variables:

(1) operator variables such as strength, body size, weight, and sex,

(2) task variables such as the coordinates of the lift, frequency of lift, etc.,

(3) object variables such as size, mass, handle details, centre of gravity, etc.

714. Approaches to modeling lifting performance differ in the extent to which they include these factors as variables. The differences reflect a trade-off of simplicity and generality against complexity and specificity. As might be expected, the different approaches can produce differing estimates of a specific operator's ability to perform a specific lifting task, reflecting the extent to which each model generalises the input variables.

715. Most manual materials handling models which have been used in practical applications deal with simple, single lifts, based on the computation of static forces. The development of dynamic models, and of models of asymmetric lifts, has proven considerably more difficult due to the sensitivity of the models to small differences in the lift characteristics. Evans (1989) reviews a three-dimensional model based on the University of Michigan Static Strength Prediction Model (section 6.3.2), and Norman and McGill (1989) report the development of a 3-D static model and a dynamic model of the lower back, based on their WATBAK model (section 6.3.4).

716. In a review of models for predicting lifting capacity, Ayoub, Mital, Asfour and Bethea (1980) categorised them into:

(1) capacity models, which use worker characteristics, task characteristics, and environmental characteristics to predict the capacity to lift; and,

(2) biomechanical stress models, which use Newtonian mechanics to estimate the stresses imposed on the musculoskeletal system of the worker during lifting.

717. The class of capacity models was divided into those based on psychophysical studies (i.e., experiments which determined the characteristics of a load which subjects were willing to lift in specific circumstances) and physiologically-based models (i.e., models based on measurements of oxygen uptake, or other physiological indices). The latter class would include the norms for lifting developed using inter-abdominal pressure as the index of effort (University of Surrey, 1980), although the norms are not models, in the sense used in this report.

718. Both classes of model have their limitations. Capacity models are typically based on regression analyses, and therefore require a large range of lifts to be studied if they are to represent a specific lift accurately. Biomechanical stress models include the details of a specific lift, but must make simplifying assumptions about the individual operator's capability. The lifting guidelines developed by the US National Institute for Occupational Safety and Health (NIOSH, section 6.3.1) are based on both psychophysical and biomechanical studies of lifting, with metabolic correction factors for repetitive lifts.

719. Current developments are towards a more integrated approach to the design of the work place combining the biomechanical models reviewed in this section with the anthropometrical models reviewed in section 6.6. Evans (1989) reviews one such development, and Kroemer et al. (1988) review the question of integrated ergonomic models in detail.

6.2.1 References

Ayoub, M.M., Mital, A., Asfour, S.S., Bethea N.J. (1980). Review, evaluation, and comparison of models for predicting lifting capacity. Human Factors, 22, (3) 257-269.

Chaffin, D.B., (1985). Computerized models for occupational biomechanics. J. Wartenweiler Memorial Lecture, International Society of Biomechanics, UMEA, Sweden.

Coerman, R.R., Magid, E.B. & Lange, K.O. (1962). Human problems under vibration stress. Human Factors, 4 (5) 315-324.

Drury, C.G. (Ed.) (1976). Safety in manual materials handling: Report on international symposium. Cincinnati, Ohio: US Department of Health, Education, and Welfare.

Evans, S.M. (1989). Use of biomechanical static strength models in workspace design. In G.R. McMillan, D. Beevis, E.Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

King, A.I. & Marras, W.S. (1988). Contribution to NRC workshop on ergonomic models of anthropometry, biomechanics, and operator-equipment interfaces. In K.H. Kroemer, S.H. Snook, S.K. Meadows, & S. Deutsch, (Eds). Ergonomic models of anthropometry, biomechanics, and operator-equipment interfaces: Proceedings of a workshop. Washington D.C: National Academy Press.

Kroemer, K.H., Snook, S.H., Meadows, S.K. & Deutsch, S. (Eds). (1988). Ergonomic models of anthropometry, biomechanics, and operator-equipment interfaces: Proceedings of a workshop. Washington D.C: National Academy Press.

University of Surrey, Materials Handling Research Unit. (1980). Force limits in manual work. (for Secretariat of Community Ergonomics Action, Luxembourg). Guildford, Surrey: IPC Science and Technology Press.

Mital, A. (1983). Preface to special issue on manual materials handling. Human Factors, 25, (5) 471-472.

Norman, R.W., & McGill, S.M. (1989). Development of objective methods for evaluating the safety of lifting tasks - biomechanical modeling. (Final report to DCIEM Toronto). Waterloo, Ontario: University of Waterloo Research Institute.

Warrington-Kearsley, Maj. B.P. (1986). A review of the NDMC back care education programme. In Backache and Back Discomfort. (AGARD-CP-378, 23-1 - 23-12), Neuilly-sur-Seine, France: Advisory Group for Aerospace Research and Development.

6.3 MODEL SUMMARIES: BIOMECHANICAL MODELS

6.3.1 National Institute for Occupational Safety and Health (NIOSH) Lifting Model

Summary Description

720. The NIOSH model is intended to predict the safe lift characteristics of tasks for a male and female civilian, industrial population. The model is based on biomechanical and psychophysical studies of lifting and includes correction factors for the shape of the load, for repetitive lifts and task duration.

721. The biomechanical aspect of the model relates the horizontal and vertical coordinates of the lift to the lumbo-sacral compressive forces at the L5/S1 intervertebral disc, and to a distance factor based on the vertical lift of the load.The metabolic aspect of the model relates the permissible load to the frequency of lifting (per minute) and the duration of lifting (hours per day), based on 33% of the assumed aerobic capacity, and on the initial posture of the lift.

722. The model takes into account:

(1) the horizontal location of the hands at the beginning and end of the lift,

(2) the vertical location of the hands at the beginning and end of the lift,

(3) the distance the load is moved vertically,

(4) the lifting frequency,

(5) the task duration.

History and Source

723. Overexertion in the work place has accounted for a large number of disabling injuries in the industrial sector (NIOSH 1981). Most of these injuries involve the act of manually handling materials, e.g. lifting. In recent years, the number of these injuries has continued to rise despite the establishment of various guidelines for such activities. Other concerns include the increasing proportion of females employed in jobs which require heavy lifting.

724. The National Institute for Occupational Safety and Health (NIOSH) Work Practices Guide for Manual Lifting was written in order to summarise the research and present recommendations to control the hazards associated with lifting (NIOSH, 1981). The guide includes a description of the NIOSH model, which was developed in order to recommend load limits for lifting tasks based upon easily measured parameters such as load weight, size, and location.

725. The model has also been incorporated in a number of computer programs. A more extensive program designed to run on micro-computers, which incorporates the NIOSH

model, is available from the Center for Ergonomics, University of Michigan, Ann Arbor, Michigan. The program provides a text and graphical output of the results of each analysis.

Product and Purpose

726. The output of the model is expressed in terms of two load limits for any given lift condition, for an industrial population. The Action Limit (AL) is the load which, given the assumptions of the model, can be lifted by 75% of healthy women and over 99% of healthy men. The Maximum Permissible Limit (MPL) is the load which, given the assumptions of the model, can be lifted by only 25% of healthy men and less than 1% of healthy women. The limits are expressed as lift characteristics which should not be exceeded: They are not risk probabilities.

When Used

727. Because of its flexibility, the model can be used either during design, to verify that component sizes and weights will be safe, or during system evaluation, to determine the current safety of a given system.

Procedures for Use

728. The user must define:

(1) number of persons performing the lift,

(2) object mass,

(3) initial hand height,

(4) final hand height,

(5) initial horizontal hand location,

(6) final horizontal hand location,

(7) rate of lift (per minute),

(8) duration of task (hours)

Advantages

729. One of the major advantages of the NIOSH model is its simplicity of use. All that is needed for the evaluation of lifting tasks is the information identified above. That information can be derived from the drawing board, during design. It can also be gathered during field trials. In the latter case, the model is simpler to use than others, which require the coordinates of arm and leg joints to be measured.

730. Another advantage of this model is that the effects of changes to the lifting task (i.e., size of the object or location of handles) can be analysed in order to determine the sensitivity of the AL and MPL to minor changes in the lifting task.

Limitations

731. The principal limitations of the model are that it is based on ideal industrial conditions:

(1) lifts in the sagittal plane,

(2) little sustained exertion,

(3) smooth, two-handed, symmetric lift,

(4) object of moderate width (75cm. or less),

(5) unrestricted lifting posture,

(6) good coupling (handles, shoes, floor surface)

(7) favourable ambient environment

(8) unaided lift (no mechanical aids).

732. The model is also limited in its assumption of user population. For example, practical use indicates that troops regularly lift at, or beyond, the AL. Whether this represents a greater capacity on the part of the troops, or greater risk taking on their part, is not clear at this time.

Application Examples

733. The model appears to be receiving increasing use in industry in North America. It has been used in human engineering evaluations of several pieces of field artillery, including 120mm mortars, 155mm howitzers, and ammunition resupply vehicles being examined by the Canadian Forces. The model was used to evaluate the more common manual materials handling tasks required to bring the weapons in and out of action, fire and maintain them. The model provided a basis for establishing appropriate manning levels for different tasks, and for a comparative evaluation of competing weapons. Application of the model resulted in the identification of several less obvious conclusions, e.g., that a two-man carry of two ammunition boxes is less demanding than a one-man carry of one ammunition box.

Technical Details

734. The programs available from the University of Michigan Center for Ergonomics run on IBM PC compatible, or Macintosh® computers. They require 128K random access memory and a graphics capability.

References

National Institute for Occupational Safety and Health (1981). Work practices guide for manual lifting. (DHSS (NIOSH) Publication No. 81-122). Washington, D.C.

University of Michigan, Center for Ergonomics (1986). Static strength prediction program. Ann Arbor, Michigan: University of Michigan.

Future Needs

735. The NIOSH model needs further validation of its applicability to military populations, and to tasks carried out in non-ideal conditions.

6.3.2 University of Michigan Static Strength Prediction Program™ (SSP)

Summary Description

736. The Static Strength Prediction (SSP) model was developed to predict static strength requirements and low back stresses in manual materials handling tasks. The prediction is based on calculations of the moments of force produced by the loads handled and by the mass of body segments. The segment masses and lengths are based on anthropometric norms scaled to body height and weight. The moments of force can be compared to muscle moments determined from individual or population strength test data. The resultant forces produced from the counteracting moments of muscle and mass cause vertebral disk compression which is also calculated by the model. The model incorporates intra-abdominal pressure (IAP). An increase in IAP due to contraction of the abdominal muscles was thought to help stabilise the spine and decrease pressure on the discs during the lift (Chaffin 1985).

737. Microcomputer based versions of the model are available, which require the following input:

(1) height and weight data for the person performing the lift, or 5th, 50th and 95th percentile population values,

(2) arm, torso, and leg postures,

(3) load magnitude and direction of action.

738. The model will output the following:

(1) stick figure posture description,

(2) resultant forces at joints,

(3) prediction of strength requirements at elbow, shoulder, hip, knee, and ankle,

(4) L5/S1 disc compression forces,

(5) abdominal pressure,

(6) male and female population strength capabilities.

History and Source

739. The model has been developed in various stages since 1968 (Chaffin, 1985). The initial version was a sagittal plane model that predicted the strength required with a given posture and load. A variation of the model found the posture that produced the highest strength capability for a given lift (Garg & Chaffin, 1975). A three dimensional model was introduced in the same paper, and is described by Evans (1989). Software for the 2-D and 3-D models is available from the Center for Ergonomics, University of Michigan, Ann Arbor.

Product and Purpose

740. The major objective of the SSP model is to predict the static strength required for any given lifting task and the low back stresses involved. This can be done for individuals or for the 5, 50, and 95th percentile population values. The model will also give the predicted strength capabilities for these percentiles in any given lifting task.

741. The information from the model can be used:

(1) to identify hazardous lifting situations where the load exceeds human limitations,

(2) to set criteria for the selection of personnel in terms of static strength,

(3) to predict task suitability for the general population by comparison to percentile capabilities.

When Used

742. The model could be used at the design stage of a task to ensure that loads, weights, and postures do not create a hazard. It could be used to evaluate the safety of existing tasks or to indicate the extent of modifications to work place or task required to reduce hazard, and to evaluate the effect of the modifications. It could also be used as a basis for worker selection standards based on height and weight.

Procedures for Use

743. The user must determine the mass to be lifted, the posture involved in the lift in terms of joint angles, and the height and weight of the person doing the lift, or the population percentile values. These are used as inputs to the Static Strength Prediction Program™. Since the model provides only a static analysis, the user must choose the point(s) of interest in the lift and perform the analysis for each posture.

Advantages

744. The model can be used to assess a task in two ways. It can indicate whether a given individual or segment of the population can perform a task, in terms of static strength, or it can assess the safety of the task in terms of compressive loading of the spine.

Limitations

745. Chaffin (1985) identified 3 limitations to the model:

(1) the model computes only compression forces on the spine, but lateral and torsional stresses are also common,

(2) the description of the muscles of the torso is too simplistic, (McGill, 1987, argues that a more detailed model will result in a different moment arm for the back extensor muscles, as noted below),

(3) the model does not account for visco-elastic creep response in the vertebral discs which occurs under prolonged loading.

746. Other limitations of the model have been noted in the review by McGill (1987):

(1) the moment arm for the back extensor muscles that is used in the model is too small, resulting in a large overestimation of compressive forces on the spine,

(2) static analyses of lifting may underestimate the peak reaction forces of the actual dynamic lifting situation,

(3) information in the recent literature casts some doubt about the role of IAP in reduction of disc compression,

(4) the model describes only single lifts and therefore does not account for the effects of repetitive lifting and any cumulative trauma incurred in lifting.

Application Examples

747. Principles from the SSP model were used in formation of the National Institute for Occupational Safety and Health (NIOSH) lifting guidelines and by the NASA Manned Spacecraft Center. The model has been used for job assessment in many industries ranging from steelworking to electronics component manufacturing. The job assessments have been used as a basis for worker selection and training.

Technical Details

748. The SSP program can be run on an IBM PC or on a Macintosh®. The requirements for the IBM PC are a graphics capability and at least 128K RAM, and DOS 2.0 or higher. The Macintosh must have a 512 K memory and Microsoft Basic® software.

References

Chaffin, D. (1985). Computerized models for occupational ergonomics. J. Wartenweiler Memorial Lecture, International Society of Biomechanics, UMEA, Sweden.

Evans, S.M. (1989). Use of biomechanical static strength models in workspace design. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Garg, A., & Chaffin, D. (1975). A biomechanical computerized simulation of human strength. AIIE Transactions, 7 (1), 1-15.

McGill, S., (1987). Issues in biomechanical modeling of the low back to determine the safe task. Communique, 17 (1), Human Factors Association of Canada.

University of Michigan, Center for Ergonomics (1986). Static strength prediction program™. Ann Arbor, Michigan: University of Michigan.

Future Needs

749. Given recent literature which challenges some of the anatomical and biomechanical features included in the model, further development and evaluation is justified, particularly in the application to dynamic lifting situations.

750. Evans (1989) describes the development of an integrated tool for Ergonomic Design using Graphic Evaluation (EDGE) which incorporates the SSP model, together with other aids for work space design. As indicated in the introduction, such approaches are the most promising long term developments of these models.

6.3.3 Job Severity Index (JSI)

Summary Description

751. The Job Severity Index is an estimation of the hazard involved in lifting tasks. It is expressed in the form of a ratio of Job Demand / Worker Capacity for a given set of job conditions. Job demand is determined by a calculation based on lifting exposure time, container size and weight, and lifting distance and frequency. Worker capacity is calculated via regression formulae based on the relationship between various individual parameters and lifting capacity. The regression model uses age, sex, anthropometric, strength and endurance measures to predict the acceptable weight for a given individual or segment of the population for the job in question. This weight is adjusted according to the frequency and height of lifting.

752. When the job includes more than one type of lift the JSI uses an equation that weights each lift depending on the duration of exposure. The output of the JSI is a ratio of demand to capacity which has been compared to injury statistics. It seems that the risk of injury starts to increase when the ratio exceeds 1.5

History and Source

753. The JSI was developed at Texas Technical University with the initial work being published in 1978 (Ayoub et al.). Additional descriptive and validation work was published in 1983 (Liles, Mahajan, & Ayoub) and 1984 (Liles, Deivanayagam, Ayoub, & Mahajan). The initial stage of development was the design of a set of mathematical models which predicted the maximum acceptable weight of lift. The mathematical relationship between the maximum weight a person felt they could safely lift (a psychophysical measure) and the frequency, object size, and height of lift was used as the basis for the models.

754. The next step was the determination of the mathematical relationship between the same maximum acceptable weight and the physical characteristics of subjects. These two areas of study produced, respectively, the job demand and worker capacity components of the JSI. The regression equations required for the JSI are presented in the papers referenced, but more information is required to perform the measurements needed for input to the model. The papers containing this information are not in the published literature but are referenced in Liles, Deivanayagam, Ayoub, and Mahajan (1984).

Product and Purpose

755. The major purpose of the JSI is the identification of those lifting tasks that create a risk of injury so that engineering or administrative controls can be implemented. The model consists of a series of regression formulae for:

(1) predicting demand and capacity, and

(2) calculating a weighted ratio for all the lifting tasks involved in a given job.

When Used

756. The JSI can be used to assess the hazard of existing lifting tasks, or of tasks in the design stage so that they may be modified if it is found that they create a significant hazard. Although the authors stress that job or task modification is the best approach, they do note that the JSI could be used as a basis for selecting workers who possess the lifting capacity required for a given job (Ayoub, Selan & Liles, 1983; Liles, Deivanayagam, Ayoub & Majahan, 1984).

Procedures for Use

757. The user must identify the following components for use in the regression models that predict task demand and worker capacity. It should be noted that the procedures for measuring worker strength and dynamic endurance were not explained in the published literature but that the reports containing the information are listed in Liles, Deivanayagam, Ayoub, & Mahajan (1984).

758. Task Demand Component Inputs:

(1) exposure in hours/day,

(2) maximum required weight of lift,

(3) frequency of lift,

(4) container size.

(5) height of lift initiation and termination.

759. Capacity Component Inputs:

(1) age, gender,

(2) isometric arm and back strength,

(3) shoulder height, abdominal depth,

(4) dynamic endurance.

760. These capacity parameters can be entered for a particular individual or a population percentile. The predicted demand and capacity values for each lifting task in the job are then used in the JSI equation which results in a ratio of demand to capacity. According to Ayoub et al. (1983), a JSI of less than 1.5 represents a nominal safety risk. The same reference also contains a table relating JSI to injury risk for a range of JSI ratios.

Advantages

761. The major advantages of this model, as compared to other biomechanical models, are its ability to combine the effects of several types of lifts within a job and its validation against actual injury statistics. The JSI also provides a comparison of task demand and worker or population capability as do some of the other models.

Limitations

762. The model is limited in that it only describes a lift in the sagittal plane and therefore does not deal with twisting actions which can be significant in injury. The JSI considers only lifting and does not deal with lowering actions (Liles, Mahajan, & Ayoub, 1983). The strength predictions are based only on static strength measures which may not accurately reflect the actual dynamic lifting situation. Although there is a dynamic endurance component in the model it was not defined in the literature reviewed. Unfortunately much of the development work which could better explain how the components of the model were derived is not published in readily available sources.

Application Examples

763. During the development phase this model was used in a large number of companies. One large study was done in an electronics equipment manufacturing plant where the JSI was applied and compared to injury statistics. Some of the concepts of the model related to the prediction of lifting capacity have been incorporated into the National Institute of Occupational Safety and Health (NIOSH) guidelines for lifting (section 6.3.1 - NIOSH, 1981). The JSI has been compared to the NIOSH lifting limits in terms of relationship to injury and was found to have a comparable potential for prediction (Liles, Mahajan, & Ayoub, 1983).

Technical Details

764. The regression equations required to model capacity and demand are presented in the form of tables of the coefficients for each input variable (Ayoub, Selan, & Liles, 1983; Liles, Deivanayagam, Ayoub, & Mahajan, 1984). The heights of lift initiation and termination are combined and classified into lifting ranges using a chart presented in Liles, Deivanayagam, Ayoub, and Mahajan. The JSI equation is also given in that reference. As noted earlier, the technical details of anthropometric, strength, and endurance measurements are not published in readily available sources but these sources are referenced in the papers listed here.

References

Ayoub, M.M., Bethea, N.J., Deivanayagam, S., Asfour, S.S., Bakken, G.M., Liles, D., Mital, A., & Sherif, M. (1978). Determination and modeling of lifting capacity (Final report). (DHEW (NIOSH) Grant No. 5R010H-00545002). Lubbock, TX: Texas Tech. University.

Ayoub, M.,Selan, J., & Liles, D. (1983). An ergonomics approach for the design of manual materials-handling tasks. Human Factors, 25 (5) 507-515.

Liles, D., Deivanayagam, S., Ayoub, M., & Mahajan, P. (1984). A job severity index for the evaluation and control of lifting injury. Human Factors, 26 (6) 683-693.

Liles, D., Mahajan, P., & Ayoub, M. (1983). An evaluation of two methods for the injury risk assessment of lifting jobs. Proceedings of The Human Factors Society 27th Annual Meeting. Santa Monica, CA: Human Factors Society.

National Institute of Occupational Safety and Health (1981). Work practices guide for manual lifting. (DHHS (NIOSH) Publication No. 81-122). Washington, D.C.

Future Needs

765. The model could be expanded to include dynamic strength measures, and lowering activities. Provision of a three dimensional analysis of the lift would also be useful. The major immediate need is for available documentation of the JSI in a complete form.

6.3.4 WATBAK- A Computer Model to Estimate Low Back Compression and Shear Forces.

Summary Description

766. WATBAK is a computer-based model used for the quantitative assessment of lifting tasks in terms of the stress placed on various body joints, including low back intervertebral joints. WATBAK accepts data obtained from representations of the lifting task, based on photographs, film, video-tape, electro-optical sensors, or a computer manikin, and incorporates those data into a computer model of the human body. The model, which is available in 2-D and 3-D versions, yields the reaction forces and moments of force at the elbow, shoulder, L4/L5 level on the lumbar spine, hip, knee and ankle joints, and estimates the size of the compressive and shear forces produced on the L4/L5 intervertebral disc. Those compressive and shear forces are compared with tolerance limits reported in the literature to determine the relative safety of lifts.

767. The concepts underlying the calculations used in WATBAK are those of conventional linked segment biomechanical models. These have been reported in biomechanical literature for many types of movements.

768. The model takes into account:

(1) body position used in the lift,

(2) weight of the object lifted,

(3) height, weight and sex of the person performing the lift.

History and Source

769. The model was developed by the University of Waterloo, Ontario, under contract to DCIEM, Toronto (Norman & McGill, 1989). The aim was to produce a tool that could be used quickly and relatively easily by minimally trained personnel, to analyse the strength demands and risk of injury for a wide variety of lifting tasks.

Product and Purpose

770. The model outputs numeric values of body segment parameters (mass, centre of gravity, length), ground reaction forces, reaction forces and moments acting on each segment, lumbar spine compression and shear forces, and intra-abdominal pressure. Maximal tolerance levels for lumbar compression and joint moments are quoted based on values obtained from the open literature. The final output of the model is a linked segment diagram of the subject in the position used in the calculations.

When Used

771. The model can be used in the design of operator tasks and work place layouts, to ensure that loads and postures do not create an unsafe risk. It can also be used to evaluate the safety of existing tasks and to indicate the extent of modifications required to reduce hazards.

Procedures for Use

772. To run the WATBAK model, the user must have a side-elevation representation of the lift to be analysed. The representation can be from digitised photographs or video-tape, from electro-optical sensors such as SELSPOT or WATSMART, or from a computer-generated manikin. From this representation, the user defines the x and y coordinates of the feet, ankles, knees, hips, shoulders, elbows, hands, head and L4/L5 of the person performing the lift. The accuracy and validity of the output is dependent on the accuracy of the digitisation. The user also inputs the height, weight and sex of the subject and the weight of the object.

Advantages

773. The major contribution of this model is the development of an easily usable, interactive software package. This package can be used to assess job task demands by an operator minimally trained in biomechanics and in the use of computers (Norman, 1984).

Limitations

774. The principal limitations of the model are:

(1) unaided lift (no mechanical aids),

(2) only static lifts in the sagittal plane can be analysed,

(3) good coupling (handles, shoes, floor surface)

(4) favourable ambient environment,

(5) only one person lifts can be analysed.

775. In addition, experience has shown that the model is highly dependent on the accuracy of digitisation of the joint centres, and that those centres cannot be located easily or reliably from photographs or video of persons wearing heavy clothing.

Application Examples

776. The model has been used in work place investigations of acute strength demands in industrial tasks by the Ontario Ministry of Labour, Ontario Workers' Compensation Board, Ontario Hydro, and several industrial organisations. The model has been used in the analysis of several military tasks. The program has also been used at the Universities of Dalhousie, Laurentian, and Queen's, and at the University of Waterloo to assess manual handling of materials problems in the mining industry.

Technical Details

777. The model was originally developed to run on a Hewlett-Packard 9845B micro-computer. The program was subsequently re-written, and will now run on any IBM-compatible PC with at least 512K. Graphical output of body postures can be obtained if the machine has a graphics card.

References

Norman, R.W. (1984). Development of a field method fc ntitative assessment of low back stresses. Final report. (DCIEM Contract No. 8SE83- 7). Waterloo,Ontario: University of Waterloo.

Norman, R.W. (1984). WATBAK: A computer package to estimate low back compressive and shear force and strength demands of manually lifting loads. (DCIEM Contract No. 8SE83-00017). Waterloo, Ontario: University of Waterloo.

Norman, R.W., & McGill, S.M. (1989). Development of objective methods for evaluating the safety of lifting tasks - biomechanical modeling. (Final report to DCIEM, Toronto). Waterloo, Ontario: University of Waterloo Research Institute.

6.4 INTRODUCTION: WORK SPACE DESIGN

778. Incompatibilities between the dimensions of the work space and the size and reach capability of the operator or maintainer can have a major effect on system performance. Typically such problems are addressed through the use of fitting trials using full-scale mockups. The work space concept must well established, however, before a mockup can be built. Fortunately, anthropometry is one of the most readily quantified of human capabilities, and lends itself to modeling in two or three dimensions. Anthropometrical models can be used to study human performance in terms the ability of an operator of a given physical size to work in a given space, reach specific controls and see specific displays. Such models are the human engineering tools which are most closely allied to engineering design methods, in particular CAD techniques (Merriman, Muckler, Howells, Olive, & Beevis, 1984).

779. Computer-generated anthropometrical models (referred to by Kroemer, Snook, Meadows, and Deutsch, 1988 as operator-equipment interface models) offer advantages over the more traditional drawing board and mock-up approaches to work space design. In particular they permit the representation of individual operators, and of the full variety of combinations of body segment lengths, rather than the more limited 'percentile' approach which has been used previously. Moroney and Smith (1972) concluded that the only solution to the problem of designing work spaces to accommodate the interaction between different anthropometric variables was the development of variable-sized mathematical man-models.

780. As well as permitting such an approach, anthropometrical models offer a medium for integrating other aspects of operator performance such as the kinematics of limb motion, force abilities, and body segment inertial abilities. At least three major attempts to produce integrated, structured, crew station design tools have been based on three-dimensional man-models (Ryan 1973; Kulwicki, McDaniel & Guadagna, 1987; Evans, 1989). Kroemer, Snook, Meadows, and Deutsch (1988) report on a workshop dedicated to examining the feasibility of developing such an integrated 'ergonomic' model for work station design.

6.5 OVERVIEW AND RECOMMENDED REFERENCES: WORK SPACE DESIGN

781. The original computer-generated anthropometrical model, or manikin, appears to have been FIRST MAN, a seven segment figure produced at Boeing Company, USA, in 1961 (Fetter, 1982). This development was followed by the more well known BOEMAN, produced at Boeing Company under Joint Army, Navy, Air Force (JANAIR) sponsorship in 1967-68. BOEMAN was part of a larger program entitled Cockpit Geometry Evaluation, intended to evaluate the physical compatibility of a seated crew member with any crew station. The original research program anticipated development starting with a link segment or 'stick figure' manikin, and progressing through six stages to incorporate enfleshment, digit manipulation, energy expenditure and force capability, and flexible skin interference analysis (Ryan , 1973). That development does not appear to have been implemented.

782. Other developments started modestly, and evolved. The SAMMIE model (section 6.6.3), which is now available commercially, originated in a simple stick figure representation of the upper torso, head, and arms (Bonney, 1969). COMBIMAN (section 6.6.2) has had a series of improvements, and a second model, CREW CHIEF (section 6.6.5), has been developed from that technology (McDaniel, 1989). This model is intended for the design of maintenance facilities in aircraft. It has several interesting features including the capability to call up standard maintenance operator postures, tools, and clothing, and the ability to

automatically adjust the manikin posture to complete the required task. CREW CHIEF also incorporates strength data relevant to maintenance tasks.

783. A number of other work space design man-models have been, or are being, developed in Australia, Canada, France, FRG, The Netherlands, U.K., and U.S.A. A review of computer-generated man-models (Hickey, Pierrynowski & Rothwell, 1985) identified two 2-dimensional anthropometry models, and thirteen 3-dimensional models. These range from 2-dimensional representations of single limbs to representations of several operators in a complex work space. They include modeling of limitations on joint motion, ability to model the effects of restraints such as safety harnesses, and ability to represent the view of the man-model.

784. The whole-body man-models vary in their level of complexity. At one extreme they are detailed representations of the body, with no built-in assumptions about posture or position. Such models are the computer equivalent of a dummy, or an articulated drawing board manikin. At the other extreme, some models include rules for the placement of the model relative to reference points such as the design eye position or the seat reference point, as well as rules governing the posture of the manikin. Both types of model include mathematical routines which develop the internal link lengths of the 'skeleton' from regression equations relating measurable external dimensions to the internal link lengths. They may also include another routine for enfleshment of the 'skeleton' once it has been generated. These mathematical routines are the "anthropometric" models of Kroemer, Snook, Meadows, & Deutsch, 1988.

785. A number of model developers have studied the incorporation of kinematic aspects of reach and vision. With the exception of the work of Ayoub (1974) and Deivanayagam, Ayoub and Kennedy (1974) few references have been found to such work, and no models appear to include a true kinematic capability. COMBIMAN includes the end points of paths of motion; TEMPUS (Badler, Korein, Korein, Radack, & Brotman, 1985; Kroemer, Snook, Meadows, & Deutsch, 1988) incorporates "goal directed reaches".

786. More detailed descriptions of the Articulated Total Body (ATB) model, BOEMAN, Computerized Accommodated Percentage Evaluation (CAPE), CAR, COMBIMAN, CREW CHIEF, PLAID/TEMPUS, and SAMMIE are provided in Kroemer, Snook, Meadows, & Deutsch, 1988. Applications of SAMMIE and TEMPUS are described in McMillan, Beevis, Salas, Strub, Sutton, and van Breda (1989). A comparative review of SAMMIE, COMBIMAN, CREW CHIEF, CAR, JACK, and SAFEWORK is provided by Paquette (1990). Despite the proliferation of models, few appear to be available to all potential users, and few appear to be used on a regular basis. The models reviewed herein appear to be the most relevant to weapon system design, although not all are widely available.

787. Comparisons of the available models are hindered by lack of details, and by lack of common definitions, such as the meaning of 'body segment' or 'link'. All of the models appear to have limitations. Typical problems to be anticipated in adopting such models are reviewed by Rothwell (1989). Typically the models lack:

(1) demonstrated validity,

(2) representation of limitations in movement and response of the limb positions to a change in posture,

(3) standard anthropometry data bases,

(4) representation of clothing and the effects of clothing,

(5) generality of the model,

(6) compatibility with the range of skills of potential users.

(7) biomechanical characteristics,

(8) compatibility with other CAD systems.

788. One feature lacking in all the models reviewed is the exploitation of the ideas of Moroney and Smith (1972) for the handling of the multi-variate anthropometrical data. With the exception of the work of Bittner (1974) and Hendy (1990), no-one appears to have exploited the potential of computer modeling of the statistics of the interactions of the different body segment dimensions. With the exception of CAR, every model reviewed used regression analyses to define 'percentile' manikins, despite the acknowledged limitations of the percentile concept to describe more than one anthropometric characteristic.

789. The most recent development of computer-generated anthropometrical man-models appears to be their use in animated work sequences, to evaluate the feasibility of performing a specific task in a confined space. Such animations have been used to investigate task performance in a zero-gravity environment, in advance of water tank simulations or actual operational experience (Emmett, 1986; Tice, 1987). Another interesting development is the move towards integration of such man-models, or manikins, with integrated human engineering design and development tools. Such tools are intended to support the design/development process from mission analysis through to detailed design (Kulwicki, McDaniel & Guadagna, 1987).

6.5.1 References

Ayoub, M.A. (1974). A biomechanical model for the upper extremity using optimization techniques. Human Factors, 16 (6), 585-594.

Badler, N.I., Korein, J.D., Korein, J.U., Radack, G.M., & Brotman, L.S. (1985). Positioning and animating human figures in a task-oriented environment. The Visual Computer, 1, 212-220.

Bittner, A.C. Jr. (1976). Computerized accommodated percentage evaluation: Review and prospectus. Proceedings, 6th International Ergonomics Association Congress. Santa Monica, CA: Human Factors Society. (157-164).

Bonney, M.C. (1969). A computerised model of a man and his environment. Paper presented at the annual conference of the Ergonomics Research Society, Bristol, U.K.

Deivanayagam, S., Ayoub, M.M. & Kennedy, K. (1974). Paths of movement of body members in aircraft cockpits. Proceedings of the Human Factors Society 18th Annual Meeting. Santa Monica, CA: Human Factors Society. (284-286).

Emmett, A. (1986, April). Simulation for the space age. Computer Graphics World, 43-48.

Evans, S.M. (1989). Use of biomechanical static strength models in workspace design. In G.R. McMillan, D. Beevis, E.Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Fetter, W.A. (1982). A progression of human figures simulated by computer graphics. IEEE Computer, 2 (9) 9-13.

Hendy, K.C. (1990). Aircrew/cockpit compatibility: A multivariate problem seeking a multivariate solution. In Recruiting, selection, training and military operations of female aircrew. (AGARD CP-491). Neuilly sur Seine: Advisory Group for Aerospace Research and Development. (15-1 - 15-8).

Hickey, D.T., Pierrynowski, M.R., & Rothwell, P.L. (1985). Man-modeling CAD programs. (DCIEM Contract No. 01SE.97711-4-8024). Toronto: University of Toronto.

Kroemer, K.H.E., Snook, S.H., Meadows, S.K., & Deutsch, S. (Eds.). (1988). Ergonomic models of anthropometry, human biomechanics, and operator-equipment interfaces. Washington D.C: National Academy Press.

Kulwicki, P.V., McDaniel, J.W. & Guadagna, L.M. (1987). Advanced development of a cockpit automation design support system. The Design, Development and Testing of Complex Avionics Systems. (AGARD-CP 417), Neuilly-sur-Seine, France: Advisory Group for Aeronautical Research and Development. (19-1 - 19-15).

McDaniel, J.W. (1989). Tools for ergonomic accommodation: COMBIMAN and CREW CHIEF. In W. Karwowski (Ed.) Computer-aided design applications in ergonomics and safety. London: Taylor & Francis.

McMillan, G.R., Beevis, D., Salas, D., Strub, M.H., Sutton, R. & van Breda, L. (Eds.). (1989). Applications of human performance models to system design. New York: Plenum Press.

Merriman, S.C., Muckler, F., Howells, H., Olive, B.R., & Beevis D. (1984). Proceedings: Workshop on applications of systems ergonomics to weapon system development. NATO Defence Research Group Panel on Applications of Human and Biomedical Sciences.(DS/A/DR(84)408). Brussels: NATO.

Moroney, W.F. & Smith, M.J. (1972). The use of bivariate distributions in achieving anthropometric compatibility in equipment design. Proceedings of the Sixteenth Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society. (16-23).

Paquette, S.P. (1990). Human analogue models for computer-aided design and engineering applications. (NATICK/TR/90/054). Natick, MA: U.S. Army Research, Development and Engineering Center.

Rothwell, P.L. (1989). Representation of man using CAD technology: User beware. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda, (Eds.) Applications of human performance models to system design. New York: Plenum Press.

Ryan, P.W., (1973). Results from a computerized crew station geometry evaluation method. In K.D. Cross, & J.J. McGrath (Eds.) Crew system design: An interagency conference. Santa Barbara, CA: Anacapa Sciences Inc.

Tice, S. (1987, March). Animated aerospace design. Computer Graphics World. 74-76.

6.6 MODEL SUMMARIES: WORK SPACE DESIGN

6.6.1 Crewstation Assessment of Reach (CAR)

Summary Description

790. CAR is intended to permit the percentage of a specified user population that is physically accommodated by the geometry of a given aircraft crew station to be determined. The CAR model is based on an internally created link structure which is generated from 14 anthropometric measures of the user population and link transformation equations.

791. The CAR link manikin is adjusted using anthropometric data, and compared with the constraints imposed by the crew station using a series of logical routines. The anthropometric data are either files of measurements from specific individuals, or are generated from existing population data (based on the 1964 US Navy aircrew anthropometry survey) using a Monte Carlo simulation process.

792. The model is normally used for studies of seated positions, although standing positions have been evaluation by NASA and the US National Bureau of Mines. Reach assessments can be carried out for up to 49 control positions. All assessments are controlled by mathematical algorithms, with no intervention by the user. A more complete description is provided in Kroemer, Snook, Meadows, and Deutsch (1988).

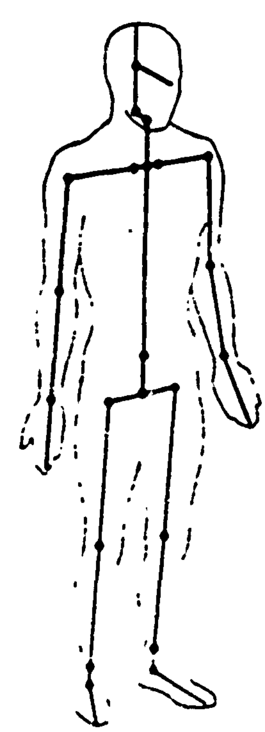


Fig. 6.1 The basic CAR link manikin

History and Source

793. The original CAR program was developed in 1976 by Boeing Aerospace Corporation for the US Naval Air Development Center (US NADC). The model was developed from the Computerized Accommodated Percentage Evaluation (CAPE) approach to using percentile anthropometry data (Bittner, 1976), and was based on the earlier BOEMAN computer-generated manikin. The impetus for CAR was the need to evaluate a number of competing candidate cockpit designs. CAR was required to be simple to operate and to require little computer time to run.

794. Under contract to US NADC, Analytics Inc. has developed and revised the program, producing versions IIA, III and IV. The latter was released in August 1984. Under a Canada:US information exchange agreement CAR IV was re-compiled to run on an IBM PC compatible computer, and evaluated for its accuracy in representing a Canadian population (Pierrynowski, 1987; Pigeau & Rothwell, 1989).

Product and Purpose

795. Each accommodation analysis produces a report summarising the results for reach, vision and head clearance tests. The contents are structured as:

(1) operator sample and crew station descriptions,

(2) percentage of operators positioned to specified anchor point,

(3) vision accommodation,

(4) percentage of operator sample accommodated,

(5) summary of reach to each control.

When Used

796. CAR was intended for use in the evaluation of proposals for crew station design. It is therefore appropriate to either the evaluation of designs as they evolve or to the comparison of different candidate concepts. CAR is also being examined for its appropriateness to assigning specific operators to crew stations which impose size limitations on the user population.

Procedures for Use

797. The user must prepare input file data for the proposed user population, and for the crew station geometry. Fourteen anthropometric variables must be specified, either for each individual operator, or for a user population. To define the user population, the user specifies: manikin anchorage point, design eye point, line of sight, seat characteristics, head clearance data, hand or foot controls. One of four anchorage points can be selected: design eye point (DEP), seat reference point (SRP), foot-point seated, foot-point standing. The crew station is defined in X,Y,Z coordinates as per normal aircraft design practice. Seat dimensions and adjustment ranges also can be specified. The user must then select the features of the crew station geometry that are to be evaluated. The program is intended for interactive use, but it

will accept data from a previously created, formatted file.

Advantages

798. CAR includes an algorithm for calculating hand reach envelopes for the operator represented by the manikin. The algorithm includes different types of grasp, and the effects of a restraint harness. This is an improvement over other models which require the individual body segments to be manipulated in order to determine the reach envelope. The ability to select one of four anchor points to which CAR will position a specific part of the link manikin, leaving the program to position the rest of the manikin appropriately, is an asset when carrying out repetitive evaluations. The use of a Monte-Carlo technique for generating the characteristics of individual 'operators' is another advantage of CAR.

799. From the technical viewpoint, CAR has the advantage of being an economical program. The original specification for CAR placed limits on the amount of CPU time, cost per run, and effort required to describe the crew station. Current versions of the model maintain the characteristics of economy in those areas.

Limitations

800. CAR's limitations arise because it is an economical program. The program has no graphical output, and the model has no enfleshment. The crew station itself is represented by X,Y,Z coordinates, which are compatible with aircraft design practice, but not with other applications. Because of the use of X,Y,Z coordinates the model cannot run 'clash routines' to indicate that the manikin is reaching through a part of the work station. The vision analysis is limited to the line of sight 'over the nose' or to the line of sight to the centre of a display from the DEP. Overall the model is a useful preliminary assessment tool, but it does not address design details.

801. Because of inherent assumptions about the crew station geometry and the visual field, CAR does not appear suited to evaluating work spaces other than aircraft crew stations (despite its use to investigate standing work situations). Evaluations of CAR III (Hickey & Rothwell, 1985) and CAR IV (Hickey & Pierrynowski, 1986) identified problems with 'user friendliness', the operator sample model, the crew station model, and the complexity of the analyses.

Application Examples

802. CAR was used to evaluate candidate crew station designs for what became the US Navy F/A-18 aircraft. It has subsequently been used to evaluate the crew stations of the H-136 Kiowa helicopter and the BAE Hawk trainer aircraft. In the latter application CAR identified some problems with the crew-station which were disproven using live subjects (Hulme & Hamilton, 1989). There have been several validation studies of the model. One approach (Bennett, Harris, & Stokes, 1982) validated 5th, 50th and 95th percentile manikins generated by CAR against the reach envelope data of Kennedy (1978). The results agreed generally within 3cm. Pierrynowski (1987) compared reach envelopes measured on 13 females with CAR predictions. In general CAR predictions underestimated the subjects' reach abilities. The results suggest that the model is too restrictive for across-body and behind-body reaches.

Technical Details

803. CAR is written in FORTRAN V, a computer language based on ANSI FORTRAN 77. The program consists of four modules. The first module allows the definition of a user sample; the second allows definition of the crew station; the third performs accommodation analyses of reach, vision and head clearance using files generated by the first and second modules; the fourth module generates reach envelopes. The CAR-IV program consists of approximately 16,000 lines of instructions, and is some 759k bytes in size.

References

Bennett, J., Harris, R., & Stokes, J.M., (1982). Crewstation Assessment of Reach (CAR) (Technical Report 1400.21B), Willow Grove, Pennsylvania: Analytics Inc.

Bittner, A.C. (1976). Computerized accommodated percentage evaluation: Review and prospectus. Proceedings of the 6th. Congress of the International Ergonomics Association. Santa Monica, CA: Human Factors Society. (157-164).

Harris, R.M., Bennett, J., & Dow, L. (1980). CAR-II: A revised model for Crewstation Assessment of Reach. (Technical Report 1400.06B), Willow Grove, Pennsylvania: Analytics Inc.

Hickey, D.T., & Rothwell, P.L. (1985). Crewstation Assessment of Reach: (CAR-III): Evaluation of the model for use in the CF aircrew/cockpit compatibility evaluation (ACCE). (Technical Communication No 86-C-04). Toronto: DCIEM.

Hickey, D.T., & Pierrynowski, M.R.(1986). Crewstation Assessment of Reach: Summary of CAR-IV revisions and reassessment for use in the CF aircrew/cockpit compatibility evaluation (ACCE). (DCIEM Contract No. 01SE.97711-4-8024). Toronto: University of Toronto, School of Physical Health and Education.

Hulme, A.J. & Hamilton, W.I. (1989). Human engineering models: A user's perspective. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Iavecchia, H., Harris, R., & Rothenheber, E. (1986). CAR IV validation study results. (Technical Report 1800.33B). Analytics Inc. Warminster, PA: US NADC.

Kroemer, K.H.E., Snook, S.H., Meadows, S.K., & Deutsch, S. (Eds.) (1988). Ergonomic models of anthropometry, human biomechanics, and operator-equipment interfaces: Proceedings of a workshop. Washington D.C: National Academy Press.

Kennedy, K.W. (1978). Reach capabilities of men and women: A three dimensional analysis. (AMRL-TR-78-50), Wright-Patterson Air Force Base, OH: USAF Aerospace Medical Research Laboratory.

Pierrynowski, M.R. (1987). Validation of sitting, torso-restrained female reach envelopes predicted by the Crewstation Assessment of Reach (CAR) model. In G. DeGroot, A.P. Hollander, P. Huijing, & G. VanIngenFschenau (Eds.). Biomechanics II. Champaign, IL: Human Kinetics.

Pigeau, R.A. & Rothwell, P.L. (1989). A validation of reach assessment in two man-modeling systems: CAR and SAMMIE. Proceedings of the Annual Conference of the Human Factors Association of Canada. Toronto. 257-260.

Future Needs

804. Representatives of US NADC have suggested that CAR might be given an interactive graphics capability, which would make it much more useful for evaluating a specific crew station. It would also permit the user to understand what the program is doing to the man-model, and to accept or reject the results on that basis. Additional improvements which are being considered include validation and improvement of the link transformation equations (focussing on female link transformations), and modelling high seat back angles, restraint systems, overhead reaches, and effects of gravity. Such developments may make the program much larger than it is currently, however, and the advantage of its economy could be lost.

6.6.2 Computerized Biomechanical Man-Model (COMBIMAN)

Summary Description

805. COMBIMAN is a three-dimensional, expert, biomechanical model of an aircraft pilot based on a link manikin. The link system is initially defined from 12 anthropometric surface dimensions; the lengths of the links can be varied to reflect different anthropometric percentiles and proportions. The user can call up data from several populations, including USAF male pilots, USAF female pilots, USAF men, USAF women, Army pilots, Army women. The user can also add additional survey data to the data base. Standard postures include sitting erect and sitting slumped against a seat back. Arm reach and reach envelopes are computed as a function of clothing and harness restraint.

806. The crew station is defined as a set of up to 250 panels, and up to 150 controls can be identified on or off the panels. Views of the model and work space can be selected by specifying the amount of roll, pitch and yaw about three orthogonal planes.

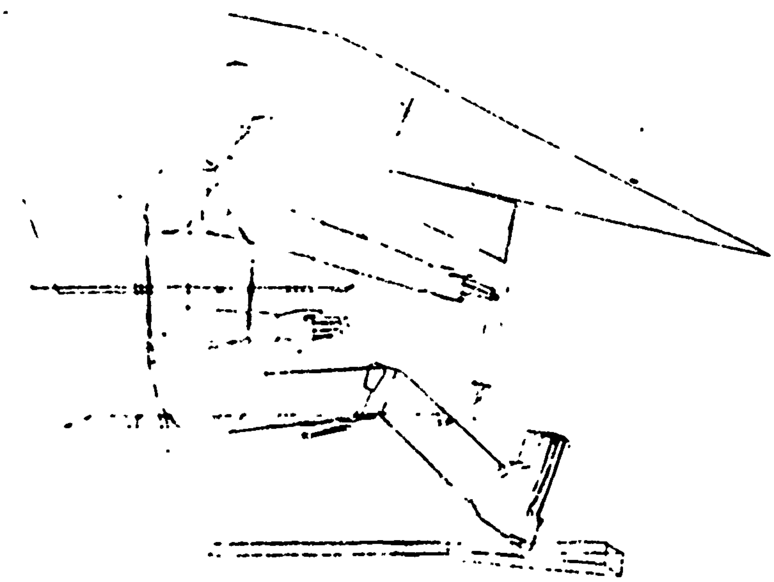


Fig. 6.2 COMBIMAN in a typical reach analysis application

History and Source

807. COMBIMAN was initially developed for the US Air Force in 1973 to assist in the design and analysis of aircraft crew stations. The program has been successively developed since that time. Version 8 was completed in 1989.

808. COMBIMAN has been distributed to major aerospace industries in the USA since 1978. The software and analytical services using COMBIMAN are available from the Crew

System Ergonomics Information Analysis Center (CSERIAC) at CSERIAC Program Office, AL/CFH/CSERIAC, Wright-Patterson AFB, OH 45433-6573, USA. US regulations limit the availability of the software outside the USA.

Product and Purpose

809. COMBIMAN produces estimates of the ability of persons with specific anthropometric characteristics to reach specified controls. COMBIMAN can also be used to produce vision plots, using the plotting algorithm centred to the eye point of the manikin. Vision plots are as per US MIL-STD-850. Strength analyses of the operation of all types of aircraft controls are also available.

810. The reach envelopes, vision plots, and strength analyses are used to evaluate proposed crew station designs for their compatibility with specific users, and thereby, with the extremes of a user population.

When Used

811. COMBIMAN, like other anthropometrical man-models, is suitable for use throughout the development of the crew station concept.

Procedures for Use

812. COMBIMAN is a task-driven, expert system. It is used interactively at a CRT terminal. The analyst or designer specifies the task COMBIMAN is to perform by answering prompts; then the program automatically simulates the activity and shows the results. The analyst selects which user population and user percentile he wishes to represent, and the 12 critical dimensions which are to be evaluated: sitting height, eye height, arm length, leg length, etc. A two-variable method of specifying the user body size range is also available. The program then provides a manikin with the most probable size and proportions, drawn from regression equations based on the USAF AAMRL Anthropometry Data Ba k. Dimensions for a select set of individual subjects may be entered to verify multivariate accommodations. The user must also specify the crew station by defining the relevant planes and controls in it. The user can then select the viewpoint from which to study the manikin and crew station.

Advantages

813. COMBIMAN is based on data drawn from an extensive anthropometric data bank. The model represents an enfleshed operator, in three-dimensional graphics. Clothing and personal equipment can also be represented, facilitating the visual examination of body clearance problems. It can perform preprogrammed reach sequences, and can simulate restraints such as shoulder harness, lap belts, etc.. Three different hand grips can be used in reach analyses (whole hand reach, functional reach, and finger tip reach). For field-of-view evaluations it is possible to add obscuration templates for helmets, respirators etc., and to represent a range of head and neck movements.

814. In addition to the ability to call up a plot of the manikin and crew station, and a visual plot from within the work space, the user can call up a zoom feature, to take a closer look at specific portions of the crew station.

Limitations

815. Because of its intended use, COMBIMAN is limited to the representation of a seated operator. US policy on technology transfer limits its availability outside the USA.

Application Examples

816. COMBIMAN has been used by the US Air Force to evaluate design changes to aircrew stations, saving the costs associated with hardware mockups and prototypes (McDaniel & Hofmann, 1990). The USAF has provided COMBIMAN software to several aerospace contractors for use in aircraft development projects. Specific details are not available.

Technical Details

817. The COMBIMAN software was written in FORTRAN IV and compiled using an IBM FORTRAN G compiler, and one module is written in IBM assembler language. The program has been re-compiled to run on the VAX series of computers, and firmware versions of the program are being distributed by CSERIAC.

References

Korna, M., Krauskopf, P., Quinn, J., Berlin, R., Rothey, J., Stump, W., Gibbons, L., & McDaniel, J. (1989). User's guide for COMBIMAN programs (Computerized Biomechanical Man-model), version 8. Wright-Patterson Air Force Base, OH: USAF Aerospace Medical Research Laboratory.

Kroemer, K.H.E. (1973). COMBIMAN - COMputerized BIomechanical MAN-model. (AMRL-TR-72-16), Wright-Patterson Air Force Base, OH: USAF Aerospace Medical Research Laboratory.

Kroemer, K.H.E., Snook, S.H., Meadows, S.K., & Deutsch, S. (Eds.) (1988). Ergonomic models of anthropometry, human biomechanics, and operator-equipment interfaces: Proceedings of a workshop Washington D.C: National Academy Press.

Hickey, D.T., Pierrynowski, M.R., & Rothwell, P.L., (1985). Man-modelling CAD programs for work space evaluations. (DCIEM Contract No.01SE.97711-4-8024). Toronto: DCIEM.

McDaniel, J.W. (1976). Computerized Biomechanical Man-Model (AMRL-TR-76-30). Dayton, OH: USAF Aerospace Medical Research Laboratory.

McDaniel, J.W. (1984) CREW CHIEF: Techniques for maintenance and workplace evaluation. Paper presented at the NATO Defence Research Group, Panel VIII workshop on Applications of Systems Ergonomics to Weapon System Development. Royal Military College of Science, Shrivenham, U.K. 9-13 April

McDaniel, J.W., Hofmann, M.A. (1990). Computer-aided ergonomic design tools. In H.R. Booher (Ed.). MANPRINT: An approach to systems integration. New York: Van Nostrand Reinhold.(205-235).

Future Needs

818. As with other models which represent a three-dimensional manikin, there is still a need to implement constraints on the interrelationships of the movements of different body segments.

819. COMBIMAN has been developed to include strength characteristics, and to compensate for the effects of clothing and harness on reach capability. Additional developments of those capabilities are planned.

6.6.3 System for Aiding Man-Machine Interaction Evaluation (SAMMIE)

Summary Description

820. SAMMIE is a commercially available computer-based anthropometrical model. The manikin is based on a link model which is enfleshed by a series of rectangular prisms or polyhedra derived from X,Y segment girth data. The link segments and other reference points are derived from anthropometric surface dimensions or from body segment lengths. The anthropometric data on which the manikin is based are derived from those reported by Dreyfus, and the girths from an RAF anthropometry survey. Other data can also be input by the user, and the Industrial Engineering Department, State University of New York has prepared a manual for data file construction from any data source.

821. The work space representation is defined three-dimensionally as basic geometric shapes (prisms, cuboids and cylinders), or by irregular solids described by vertices, edges and faces. The latest version of the program provides 'clash' detection, to identify if work space entities interfere with one another, and by how much. The program also includes a surface and shading facility to produce realistic views of the model and work space.

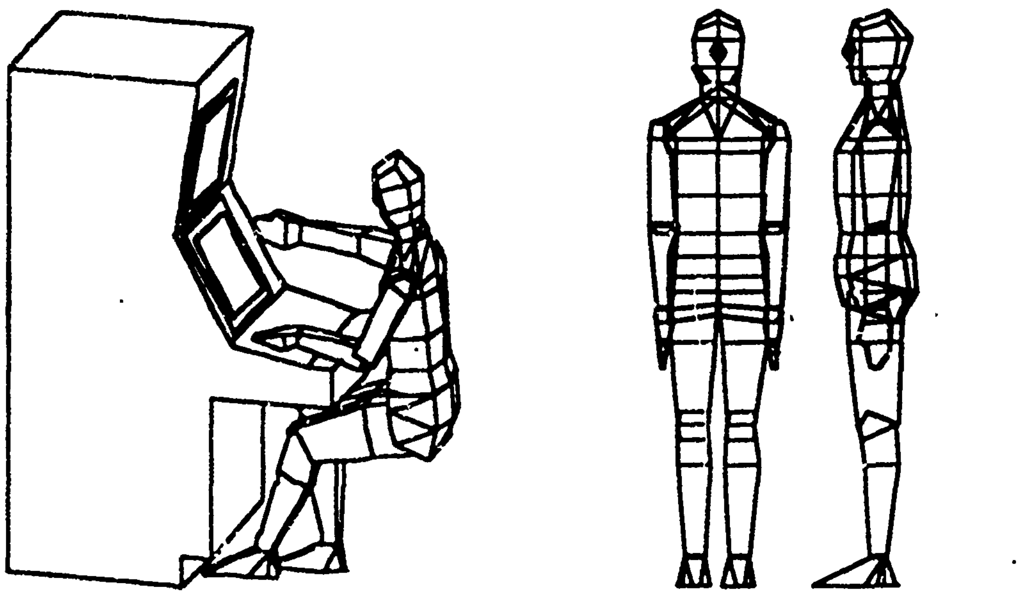


Fig. 6.3 Typical application of SAMMIE, and examples of basic manikin

822. Work space details can be generated interactively, or can be entered to a data file following off-line preparation. The spatial and hierarchical relationships of the components of the work space must be specified. Because of those relationships, mechanical functions can be simulated, such as the upward and downward movement of the forks on a fork lift truck. Such movements may be grouped (all members of a set move together), or independent.

History and Source

823. The development of SAMMIE is reported by Bonney (1989). SAMMIE was developed at the University of Nottingham, UK, in 1968/69. The original version was a stick figure upper torso. It was developed to be a general purpose manikin operable in a general purpose work space, and marketed in the UK by Compeda Inc. In 1982 the program rights were purchased by Prime Computer Limited, and the program was improved for commercial distribution. It became available for industrial use in North America in 1984, running on a Prime® computer. In 1989 Prime discontinued full support of the software. In 1987 the original developers of the model launched a separate version of the model, SAMMIE C.A.D, which runs on Sun®, VAX® and Silicon Graphics Iris® computers, and is marketed worldwide by SAMMIE C.A.D, Loughborough, U.K..

Product and Purpose

824. SAMMIE has available a variety of graphical representations, including mirror and reflection views, three orthogonal views and perspective view, and mesh-grid field-of-view representation. Such views can be selected either from the manikin's eye position or from an external reference point. The program also permits the user to test reach and sight to a sequence of specific points in the work place.

When Used

825. As with other computer manikins, SAMMIE can be used throughout the design/ development cycle, to address questions of work space design.

Procedures for Use

826. SAMMIE is used interactively from a graphics terminal. The program is structured in modules that address different characteristics of the model: creation of primitives to build work place models, creation of the model structure, control of the 3-D view of the work place, storage/retrieval of models, production of Aitoff or Mercator projections, removal of hidden lines, field of vision, movement of the man-model's limbs, control of the man-model's anthropometric characteristics, etc. The user accesses each of these modules using the keyboard, tablet or screen menu. Several analytical functions can be completed in 'background' while the user makes up the viewing features on the CRT.

Advantages

827. SAMMIE is a general purpose representation of operators and their work space. The program permits the representation of multiple operators, and elaborate visual fields. The hierarchical construction of the work space representation permits it to be manipulated to simulate the movement of equipment, for example the movement of objects seen from within a vehicle. The manikin includes limits for joint rotation, including 'comfort' limits.

Limitations

828. A major limitation is that SAMMIE currently does not incorporate anthropometry data from representative populations. In addition, some of the fixed characteristics of the link model result in unrepresentative values of manikin dimensions when representing personnel

near the tails of the population size distribution.

829. As with other computer manikins SAMMIE does not incorporate any logic to modify posture as a result, for example, of changing the line of sight.

Application Examples

830. SAMMIE has been used in a variety of design projects. Bonney (1989) reports that the applications fall into two broad categories. The first is the design of crew stations for aircraft, vehicles, ships, etc. The second in the design of computer-based work stations such as CAD terminals, bank teller stations etc. More than thirty organisations in the U.K., and at least six companies in the USA have used the program. SAMMIE has been used by the Canadian Forces for the evaluation of existing aircrew size selection standards (Rothwell, 1989).

Technical Details

831. The SAMMIE program marketed by Prime Computer Ltd. runs on all Prime 50 series 32-bit computers. The program is written in FORTRAN 77 and has the facility for 2-way transfer of models created using other Prime CAD programs. The SAMMIE program marketed by Prime Computers runs on all Prime 50 series 32-bit computers. The program is written in FORTRAN 77 and has the facility for 2-way transfer of models created using other Prime CAD programs. Details of the SAMMIE C.A.D. version are not available. It runs on a VAX computer, Sun work station, or Silicon Graphics Iris.

References

Binder, L. (1988). "SAMMIE" crew station design. Poster presentation to NATO workshop on Applications of Human Performance Models to Systems Design. St. Louis, MO: McDonnell Aircraft Co.

Bonney, M. (1989) Applications of SAMMIE and the development of man-modelling. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L.van Breda (Eds.). Applications of human performance models to system design. New York: Plenum Press.

Bonney, M.C., Blunsden, C.A., Case, K., & Porter, J.M. (1979). Man-machine interactions in work systems. International Journal for Production Research. 17 (6), (619-629).

Bonney, M.C., Case, K., & Porter, J.M. (1982). User needs in computerised man models. In R. Easterby, K.H.E.Kroemer (Eds.) Anthropometry and biomechanics: Theory and practice. New York, Plenum Press, (97-101).

Case. K., & Porter, M. (1980, January). SAMMIE. A computer-aided ergonomic design system. Engineering.

Case, K., & Porter, M. (1980). SAMMIE can cut out the prototypes in ergonomic design. Control and Instrumentation. (28-29).

Kingsley, E. C., Schofield, N.A., & Case, K. (1981). SAMMIE. A computer aid for man machine modelling. Computer Graphics, 15 (3), 163-169.

Porter, J.M., Stearn, M.C., Geyer, T. A.W., Smith, P.A., & Ashley, R.C. (1984). An evaluation of the usefulness of SAMMIE in vehicle design. In E.D. Megaw (Ed.). Contemporary ergonomics. London: Taylor and Francis, 272-276.

Rothwell, P.L. (1989). Representation of man using CAD technology: User beware. In G.R. McMillan, D. Beevis, E. Salas, M.H. Strub, R. Sutton, & L. van Breda, (Eds.) Applications of human performance models to system design. New York: Plenum Press.

Rothwell, P.L., Pigeau, R.A. (1990). Anthropometric accommodation of females in Canadian Forces aircraft crew stations. Recruiting, selection, training and military operations of female aircrew. (AGARD-CP-491). Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development.

Future Needs

832. Several developments of SAMMIE are currently underway, including the development of the US military standard anthropometry data files. The most desirable development is seen as the provision of the equivalent of a reflex system, to control the posture of the manikin as individual body segments are manipulated. A second logical extension would be to include biomechanical modeling, so that the manikin can be used to evaluate manual materials handling tasks.

6.6.4 Anthropometric Design Assessment Program System (ADAPS)

Summary Description

833. ADAPS is an interactive graphical computer-aided design tool for visualisation of a human work space with an emphasis on anthropometrical evaluation. With ADAPS, three-dimensional anthropometrical models, based on 24-element link systems, can be displayed on a graphics screen in a three-dimensional drawing of a work space.

834. Enfleshment is provided by straight lines drawn between link-related surface points. Link segments and surface points are derived from anthropometric surface dimensions or body segment lengths. The anthropometric data on which the manikins are based are derived from those reported by Molenbroek and Dirken (1987). Other data can be input by the user, to represent specific populations such as the elderly or handicapped.

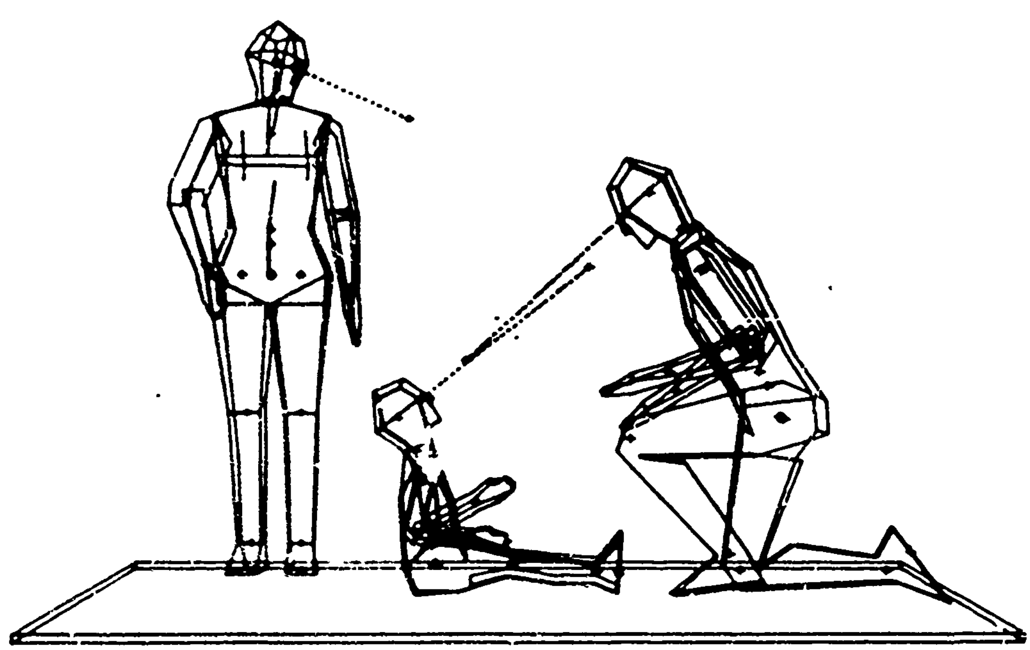


Fig. 6.4 Representation of some of the ADAPS-manikins

835. Features of ADAPS include:

(1) interactive manipulation of the manikin's posture with automatic restriction of joint-range angles,

(2) interactive definition of body size to represent the desired population and size category (percentile),

(3) integrated reach algorithms for the hands and feet and a direction algorithm for the manikin's field of view.

836. Work space representation is defined as three-dimensional wire frames, using basic geometric shapes (cuboids and cylinders), translation and rotation sweeps, or plain lines. The work space description has to be entered off-line, before storage in a 'work space library'. After this storage the user can switch between design alternatives in a matter of seconds. Hierarchical relationships of the work space components can be specified in such a way that functionality can be simulated.

History and Source

837. ADAPS was started in 1979 as a graduate project in the Product Ergonomics Group of the Faculty of Industrial Design Engineering, Delft University of Technology (DUT), The Netherlands. Further development, in cooperation with the Computer Center of DUT, resulted in a practical and efficient CAD tool for work space design and anthropometrical assessment.

Product and Purpose

838. ADAPS can produce orthogonal, isometric, and perspective views of the work space and the manikin. It can also produce views from the manikin's eye point. The program therefore permits examination of fit, reach, access, and view to specific points in the work space. The main purpose envisaged for ADAPS at present is in research and education of industrial designers. ADAPS can be used throughout the design process, but is seen as particularly relevant to the early stages, when different work space design options can be evaluated. In teaching, it is currently used in the curriculum of the Faculty of Industrial Design Engineering by means of:

(1) short (three afternoons) introductory practice sessions in anthropometrical assessment, twice a year, during the student's second year,

(2) 40 hour courses in computer-aided design and anthropometrical assessment,

(3) an assessment tool in design or graduate projects which lead to development of a working prototype.

Procedures for use

839. ADAPS is used interactively from a standard computer terminal and a number of graphic display screens. The modular structure of the program permits different characteristics to be controlled through the keyboard, screen menus, or control dials. The user must define the work space using the geometric primitives, define the user as a percentile of the 'standard' population, or enter data for a specific user population.

840. Real-time control of the manikin's posture is possible, or postures can be stored and retrieved. The manikin's limbs are controlled either by the input of joint angles or by using reach or displacement algorithms. The field of view from the manikin's eye point can also be controlled.

Advantages

841. ADAPS is a fast, simple, and easy to learn computer work space design tool, with an emphasis on anthropometrical assessment of work space designs. The program permits the representation of multiple manikins. Manipulation of posture is claimed to be 'user friendly' because of the reach and displacement algorithms.

Limitations

842. A major limitation of ADAPS is that it is not, strictly, available commercially. It runs only on the PDP 11-44 series of computers, and uses a graphic language that is becoming outdated. Surface-modelling and solids-modelling are not possible. Creation of the workspace cannot be specified or manipulated interactively, and definitions of mirrors or reflecting surfaces are not possible.

Application Examples

843. Although not a commercial product, ADAPS has been sold to a number of institutions in The Netherlands, and a version has been implemented in a car manufacturing company. Some of the industrial uses include the design and evaluation of a check-out counter and car design.

844. Graduate student projects in which it has been used include:

(1) the design of a work table for a wheelchair occupant,

(2) accessibility for inspection of an automated dairy,

(3) assessment and redesign of a car for the handicapped,

(4) design and evaluation of a magnetic pay-card reader placed in city and regional buses and trams.

Technical Details

845. ADAPS runs only on the PDP 11-44 series of computers, under the RSX operating system, using GPGS graphics software. The program is written in FORTRAN. Graphics

output is on a Vector General series 3 vector refresh display, a Tektronix 4010-compatible storage display, or a colour raster scan display.

References

Beimers, H.W. (1987). Een 3-D mannikin voor Volvo Car b.v. Delft. The Netherlands: Delft University of Technology, Faculty of Industrial design Engineering.

Hoekstra, P.N. (1985). ADAPS - Herkomst van de gegevens van het antropometrisch model. Delft. The Netherlands: Delft University, Faculty of Design Engineering.

Hoekstra, P.N. (1987). Education and research in computer-aided work place assessment. Proceedings of INTERFACE 87: Human Implications of Product Design. Rochester, New York.

Molenbroek, J.F.M., and Dirken, J.M. (1987). Nederlands lichaamsmaten voor ontwerpen. DINED-tabel Delft. The Netherlands: Delft University of Technology, Faculty of Industrial Design Engineering.

Future Needs

846. Several developments of ADAPS are currently underway:

(1) the implementation of ADAPS on VAX computers using GKS (Graphical Kernel System),

(2) the implementation of an ADAPS version on PC level computers,

(3) the evaluation and extension of an experimental biostatic force model of ADAPS,

(4) the evaluation and extension of the anthropometrical data base for automatic manikin generation.

6.6.5 CREW CHIEF - A 3-D Computer Model of a Maintenance Technician

Summary Description

847. The 3-D CREW CHIEF model provides the designer with the ability to simulate, on the computer-aided drawing board, maintenance and other related human operator interactions with a system. It creates human models in the size range 1st to 99th percentile for both male and female maintenance technicians, the encumbrance of four types of clothing and personnel protective equipment (PPR), joint mobility limitations which are a function of clothing, 12 working postures, automated physical accessibility for reaching into confined areas (with hands, 105 hand tools, and objects), visual access (evaluating what the CREW CHIEF can see), strength capability (for using wrenches and manual material handling tasks). It is claimed that CREW CHIEF is an expert system which enables the designer to perform the functions of an expert ergonomist.

History and Source

848. CREW CHIEF is a joint development of the U.S. Air Force's Armstrong Aerospace Medical Research Laboratory and the Human Resources Laboratory at Wright-Patterson AFB, Ohio. Version 2 of the software was completed in 1989 with ability to add new tools to the data base, add new body size surveys to the date base, and display shaded surface models. The software or analytical services using CREW CHIEF are available form the Crew System Ergonomics Information Analysis Center at CSERIAC Program Office, AL/CFH/HE/CSERIAC, Wright-Patterson AFB, Ohio, OH 45433-6573, USA.

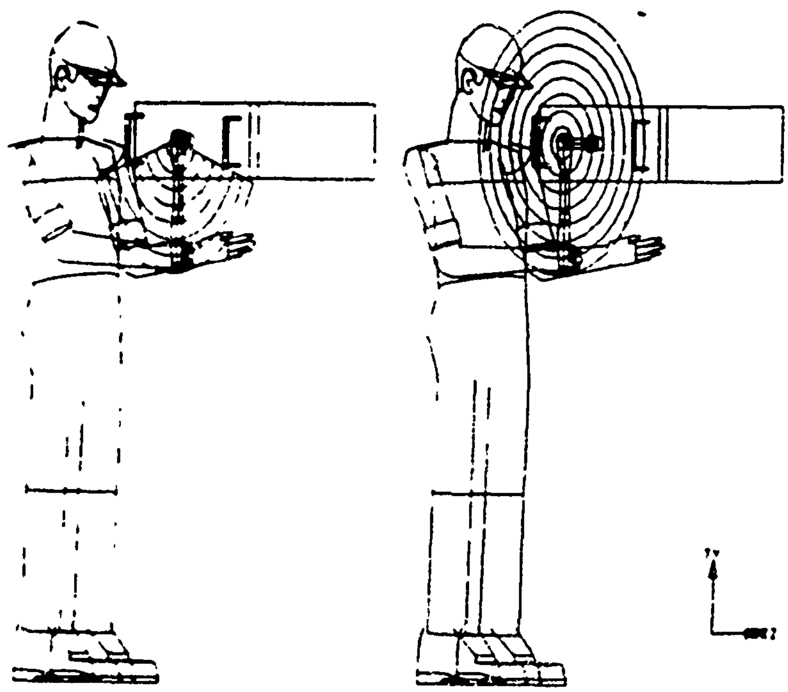


Fig. 6.5 Typical application of CREW CHIEF: left figure shows envelope of ratchet tool interferes with handles on a box; right hand figure shows how extension of ratchet socket permits unobstructed use of tool.

Product and Purpose

849. Early identification of potential design-induced maintainability problems is essential to correct problems before mock-up, fabrication, or production. CREW CHIEF simulates a maintenance activity on a computer-generated image of the proposed system design, to determine if the activity is feasible. Since it incorporates sophisticated models of extensive ergonomics data, the designer can use the model without having to be an expert ergonomist.

850. Approximately 35 percent of the lifetime equipment cost, and one-third of all manpower, is spent on maintenance. Excessive repair time is caused by failure to consider maintenance demands adequately. Maintenance technicians can spend hours making repairs which could be completed in minutes with better accessibility. CREW CHIEF is intended to reduce the incidence of such problems by allowing the designer to perform maintainability analyses and correct design-related defects. Ultimately, not only will development costs and acquisition time be reduced, but also life cycle costs and maintenance time will be reduced and systems availability increased.

When Used

851. This model, interfaced to existing commercial CAD systems used by aerospace manufacturers, may assist in evaluating the maintainability of aircraft, and of equipment in general. The CREW CHIEF model allows the designer to simulate a maintenance activity using the computer-generated design. The need for details of the physical design, replaceable units, etc. makes the model most suitable for use in the preliminary design and detailed design stages.

Procedures for Use

852. The user answers a series of questions which define the maintenance task to be simulated. Most answers are selected from menus. On-line 'help' explains the choices. Once the task is defined, the analysis is automatic.

853. The designer may simulate a maintenance activity on the computer-generated system, to determine if it is feasible. Expert system software automatically creates the correct body size and proportions for males and females, the encumbrance of clothing, personnel protective equipment, and mobility Physical access for reaching into confined areas (with hands, tools, and objects), visual access, and strength analyses are conducted. At the conclusion of the analysis, the 3-D human model is displayed, superimposed on the design, performing the task under analysis

Advantages

854. Since CREW CHIEF is interfaced to existing commercial CAD systems used by aerospace manufacturers, the program does not require users to enter the design into the CREW CHIEF program for analysis rather CREW CHIEF is called into the user's design 'drawing' without any conversion. CREW CHIEF operates as a subprogram to the CAD system, and is always readily accessible It is a task-driven expert system. The user need only understand the task to be simulated, without needing expert knowledge of ergonomics.

Limitations

855. CREW CHIEF is a simulation of empirical data, and is limited to those types of analyses covered by the available data bases. Because it is hosted with only a few CAD systems, it is not available to everyone, although the analysis services are available through CSERIAC.

Application Examples

856. CREW CHIEF has been used in numerous applications for analysing accessibility with hand tools in environments with limited accessibility (McDaniel & Hofmann, 1990). The capability to analyse strength required for materials handling in unusual maintenance postures has been used in many non-military projects, to verify performance.

Technical Details

857. CREW CHIEF is currently interfaced to the following:

(1) CREW CHIEF Host-independent, (unhosted core of CREW CHIEF), FORTRAN 66 and FORTRAN 77.

(2) CREW CHIEF - CADAM Version 20, with Geometry Interface Module (GIM) for MVS/SP operating system, FORTRAN 66H.

(3) CREW CHIEF - CADAM Version 21, with Interface User Exit (IUE) for MVS/SP operating system, FORTRAN 66 and 77, and for VM/IS operating system, FORTRAN 77H Extended.

(4) CREW CHIEF - Computervision Version CADS 4001, with Analytical Processing Unit (APU) and CADDS 4X software, revision 5B or later.

(5) CREW CHIEF - Computervision CADDStation Version for UNIX operating system.

References

Ayoub, M.M., Smith, J.L., McDaniel, J.W., & Selan, J.L. (1985). Job demands in some military maintenance operations. Proceedings of the 9th Congress of the International Ergonomics Association. Bournemouth, England.

Korna, M., Rothey, L., Jones, M., Krauskopf, P., Stump, W., Hardyal, S., Haddox, D., Meeks, L., & McDaniel, J.W. (1988). User's guide for CREW CHIEF: A computer graphics simulation of an aircraft maintenance technician (Version 1-CV4001). (AAMRL-TR-88-045), Wright-Patterson AFB, Ohio. Armstrong Aerospace Medical Research Laboratory.

Korna, M., Krauskopf, P., Haddox, D., Hardyal, S., Jones, M., Polzinetti, J., & McDaniel, J.W. (1988). Users guide for CREW CHIEF: A computer graphics simulation of an aircraft maintenance technician (Version 1-CD20). (AAMRL-TR-88-034). Wright-Patterson AFB, Ohio: Armstrong Aerospace Medical Research Laboratory.

McDaniel, J.W. (1984). CREW CHIEF: Techniques for maintenance and workplace evaluation. Paper presented at the NATO Defence Research Group, Panel VIII workshop on Applications of Systems Ergonomics to Weapon System Development. Royal Military College of Science, Shrivenham, U.K.

McDaniel, J.W. (1985). Computer-aided design models to support ergonomics. Proceedings of the 9th Congress of the International Ergonomics Association, Bournemouth, England..

McDaniel, J.W. (1989). Tools for ergonomic accommodation: COMBIMAN and CREW CHIEF. In W. Karwowski (Ed.) Computer-aided design applications in ergonomics and safety. London: Taylor & Francis.

McDaniel, J.W. (1989). Modeling strength data for CREW CHIEF. Proceedings of the SOAR 89 (Space Operations, Automation, and Robotics), Houston, TX: NASA Johnson Space Center.

McDaniel, J.W. & Askren, W. (1985). Computer-aided design models to support ergonomics. (AAMRL-TR-85-075). Wright-Patterson AFB, Ohio: Armstrong Aerospace Medical Research Laboratory.

McDaniel, J.W., Hofmann, M.A. (1990). Computer-aided ergonomic design tools. In H.R. Booher (Ed.). MANPRINT: An approach to systems integration. New York: Van Nostrand Reinhold. (205-235).

CHAPTER 7

TRAINING AND SKILL RETENTION MODELS

7.1 INTRODUCTION

858. Over the past few years, researchers and practitioners from many disciplines (e.g., engineering, education, psychology) have increasingly depended on more complex methodologies, techniques and procedures for the assessment of complex human behaviour. Indeed, research and development (R&D) has intensified since there is an imperative need to make more effective use of human performance data. In response to this need, much of the human performance research and applications have focused on developing uniform concepts, definitions, categories and measures to allow better utilisation and generalisation of research findings to operational environments such as military training (e.g., Gagne, 1965; Fleishman, 1982; Levine, Romashko and Fleishman, 1973; Vreuls and Obermayer, 1985).

859. The military, which devotes considerable effort and resources to the enhancement of training systems, has benefited from the aforementioned R&D efforts. Contributions to the military training include specifying ability requirements for certain tasks, deriving performance-task taxonomies and feedback systems, aiding in design decisions for man-machine systems, and developing models of human performance for training management (i.e., what and how to train, which skills are easy or difficult to learn or retain, etc.) (see Fleishman, 1975; Peterson and Bownans, 1982; McCormick, 1976; Wickens, 1984). However, the increased sophistication and complexity of current and emerging weapon systems and training systems makes critical further development of human performance models (especially those that deal with training and skill retention) for analysing, designing and evaluating training systems.

860. The purpose of this section is twofold: (1) to outline and briefly describe cognitive, mathematical, task- based, and system-oriented models of training and skill retention available in the literature and (2) to highlight the important applications of these models to military training system design.

7.2 OVERVIEW AND RECOMMENDED REFERENCES

861. In order to define the domain under review and to establish guidelines for the review procedure, a definition of the term "model" in the context to training was needed. The definition used was adapted from Meister (1985):

> A model of training and skill retention describes in quantitative terms (i.e., mathematics), words (i.e., set of theoretical assumption/rules), or graphical symbols (i.e., organisational framework), the cognitive and behavioural events and processes involved in learning (acquiring, retaining, and maintaining) specific task-related performance.

862. The rationale for adopting this definition was that the information derived from a preliminary literature review provided a narrow view of the training and skill retention field. That is, only models that provide quantitative predications of performance; those that could be implemented on a computer or dealt with procedural tasks were depicted. Clearly there are other models--descriptive in nature--(e.g., automatic/control processing model, cognitive

models) that do not have any quantitative characteristics but contribute significantly to the understanding of task performance in training environments. These "descriptive models" explain observed behaviour ("what humans do"), they are rule-based, and have almost no predictive capabilities. In addition, descriptive models provide the following advantages: (1) organise and synthesise research; (2) generate hypotheses for empirical testing and validation; (3) simplify complex interrelationships with real-world applications and (4) provide a useful framework for interpreting human performance data. Therefore, the inclusion of these types of models is warranted.

863. Three conclusions can be offered concerning the state of the art in this area. First, from the review of the literature it is clear that these models of training and skill acquisition need further development, refinement, and validation on a variety of tasks (not only procedural). Most of the models must expand their scope to be of use to training researchers and practitioners. Second, Sticha, Knerr and colleagues (Knerr and Sticha, 1985; Sticha, 1982; Sticha, Blacksten, Mumaw, Morrison, Deyoe, Cross, Buede, and Zirk, 1986; Sticha, Edwards, and Patterson, 1984) have generated the best research and demonstration of models of skill retention. They have applied SAINT, incorporated psychological models and theories of skill acquisition to explain the training of procedural tasks. Third, as stated before, their models are of limited use to the training manager. What is needed are more general models of training where issues such as task, skill, ability, device design, engineering and cost variables are incorporated. Such models can help to simplify complex task interrelationships on the job, making training design more straightforward, and training goals more objectively defined. Such models also may offer users a better understanding of task performance in training environments, i.e., what the trainee does in a behavioural sense.

7.2.1 References

Anderson, J. R. (1982). Acquisition of cognitive skill. Psychological Review, 89 (4), 369-406.

Atkinson, R. C., Bower, G. H., & Crothers, E. J. (1965). An introduction to mathematical learning theory. New York: John Wiley & Sons.

Fleishman, E. A. (1975). Toward a taxonomy of human performance. American Psychologist, 30, 1127-1149.

Fleishman, E. A. (1982). Systems for describing human tasks. American Psychologist, 37, 821-834.

Gagne, R. M. (1965). The conditions of learning. New York: Holt, Rinehart & Winston.

Knerr, C. M. & Sticha, P. J. (1985). Models of learning and performance of armor skills. Proceedings of the Symposium on the Military Value and Cost-Effectiveness of Training (pp. 491-572). Brussels, Belgium: NATO

Levine, J. M., Romashko, T., & Fleishman, E. A. (1973). Evaluation of an abilities classification system for integrating and generalizing findings about human performance: The vigilance area. Journal of Applied Psychology, 58, 147-149.

McCormick, E. J. (1976). Job and task analysis. In M. D. Dunnette (Ed.) Handbook of industrial and organisational psychology. Chicago: Rand McNally.

Meister, D. (1985). Behavioral analysis and measurement methods. New York: John Wiley & Sons.

Peterson, N. G., & Bownans, D. A. (1982). Skill, task structure and performance acquisition. In M. D. Dunnette & E. A. Fleishman (Eds.) Human performance and productivity, human capability assessment (Vol. 2). Hillsdale, N. J.: Lawrence Erlbaum.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. Psychological Review, 84, 1-66.

Sticha, P. J. (1982). Review of analytical models of procedural learning and performance. Fort Knox, Kentucky: U.S. Army Research Institute.

Sticha, P. J., Blacksten, H. R., Mumaw R. J., Morrison, J. E., Deyoe, P. W., Cross, K. D., Buede, D. M., & Zirk, D. A. (1986). Optimization of simulation-based training system. Alexandria, VA: Human Resources Research Organization.

Sticha, P. J., Edwards, T. D., & Patterson, J. F. (1984). An analytical model of learning and performance of armor procedures (Research note 84-24). U.S. Army Research Institute for the Behavioral Sciences, Alexandria, VA.

Vreuls, D. & Obermayer, R. W. (1985). Human system performance measurement in training systems. Human Factors, 27, 241-250.

Wickelgren, W. A. (1981). Human learning and memory. Annual Review of Psychology, 32, 21-52.

Wickens, C. D. (1984). Engineering psychology and human performance. Columbus: Charles E. Merrill Publishing Company.

7.3 COGNITIVE MODEL SUMMARIES

7.3.1 Anderson's ACT Production System of Skill Acquisition

Summary Description

864. In the ACT Production System model three processes of skill acquisition were identified by Fitts (1964) and Fitts and Posner (1967): (1) a cognitive process -- initial encoding of this skill sufficient to produce crude performance of the task; (2) an associative process -- involves "smoothing out" the skill performance; and (3) an autonomous process -- continued, but gradual, improvement of skill performance. The ACT system, a reformulation of Fitts theory, consists of three corresponding stages: (1) a declarative stage -- where learner receives instruction and information about a skill, (2) a knowledge compilation stage -- where practice of a skill converts declarative knowledge to procedural form; and (3) a procedural stage -- the turning of knowledge into specific applications.

History and Source

865. The model is based on the Fitts (1964) processes of skill acquisition theory The original theory was developed to help explain cognitive prerequisites for training transfer.

Product and Purpose

866. The purpose of this system is to improve skill acquisition by presenting the learner with a training system structure and organisation which best fits the processing strategies employed by the learner.

When Used

867. This model is best applied to develop intelligent tutoring systems or computer assisted instruction. It can be used to guide instructional developers and training device designers.

Procedures for Use

868. This model cannot be applied "as is." That is, it needs to be integrated with task requirements, instructional purpose, and system limitations.

Advantages

869. The advantages claimed for the model are: (1) it can be applied as a computer-based tutor in standard classroom instruction; and (2) the ACT Production System provides a mechanism to design and develop compatible curriculums, hence making learning easier.

Limitations

870. So far this model has been applied to simple tasks only (e.g., geometry, production systems, proof-reading).

Application Examples

(1) Development of the "Official Production System" (OPS) language for representing basic, instructable production systems.

(2) Planning activities and representing strategic knowledge in production systems.

Technical Details

871. Much of the ACT performance theory is concerned with specifying how productions are selected for application, while ACT learning theory is concerned with how these production rules are acquired. There are three unique features of ACT. In ACT: (1) strength increases linearly rather than exponentially; (2) increase of strength is only one of several mechanisms by which learning occurs, rather than a principal mechanism; (3) strength is only one of several criteria used to determine which production is actuated.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. Psychological Review, 89 (4), 369-406.

Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), Categories of human learning. New York: Academic Press.

Fitts, P. M., & Posner, M. I. (1967). Human performance. Belmont, CA: Brooks/Cole.

Future Needs

872. An immediate need is the application of the ACT Production System Model to more complex tasks, such as decision-making (i.e., less procedural).

7.3.2 Controlled and Automatic Human Information Processing Model

Summary Description

873. This model argues that there are two processes in human cognition: (1) automatic processing, which is a fast, parallel process that is unavailable to conscious awareness, not limited by Short Term Memory (STM), requires little subject effort, and requires extensive and consistent training to develop, and (2) controlled processing, which is slow, serial, conscious, limited by STM, and requires little or no training to develop.

874. The model can be explained by either a resource view, which considers performance to be automatic when processing is parallel and virtually no resources are required, or a memory view, which considers performance to be automatic when there is single step, direct access retrieval from memory. The former view is better able to explain why novice performance is so poor, while the latter better explains how automatic performance is learned (and why consistency is so important in the development of training).

History and Source

875. The model is based on Atkinson and Shiffrin's (1968) information processing theory of verbal memory. Hasher and Zacks (1979) also noted that some automatic processing is based on heredity (i.e., recording of frequency, spatial, and temporal information), but the majority of tasks become automatic only through learning. Furthermore, LaBerge and Samuels (1974) noted that acquiring automatic processes through practice (i.e., on component skills) may be necessary to learn complex skills.

Product and Purpose

876. The model has been used to explain reduction in information demand effects in visual and memory search tasks and reductions in dual task interference with practice. The model has also been used to explain automatic perceptual and motor skills such as driving and riding a bicycle and cognitive skills such as reading and visualising a triangle.

When Used

877. The model is best used to develop guidelines for training automatic skills (see Schneider and Shiffrin, 1977). Instructional developers can use the guidelines to design, develop, and implement training techniques. In addition, part-task training strategies can be derived.

Procedures for Use

878. The model cannot be applied "as is." It is more of a theory or conceptual framework than a model. However, guidelines for instructional strategies can be derived to both aid in the development of automaticity and to differentially present information based upon the type of processing taking place. That is, in controlled processing, increases in presented information requires increases in time for processing, while automatic processing does not require the added time element. Furthermore, the type of processing taking place can be determined in that automatic processes are not influenced by intentional learning, instructions or practice, the performance of concurrent tasks, or affective states. Finally, context effects

may play a role. "Practitioners may have little control over context at application, which is often determined by other factors such as the mission to be accomplished and the equipment to be used. However, he or she can control the context at training and should design it to take into account the breadth and nature of the context at application" (Logan, 1988, p. 591).

Advantages

879. The model can explain retention of skills over large periods of disuse and also workload reductions with training. A microcomputer or some form of automated training makes this kind of drill and practice feasible.

Limitations

880. The development of automatic processing requires extensive drill and practice (hundreds of training trials). Also, the model has been used to explain perceptual and motor skills, but has focused less on explaining how cognitive tasks become automatic. In addition, it is difficult to derive and to define consistent components of a whole task. Finally, the model is not quantitative in nature and reflects more of a conceptual framework than a model of human performance.

Application Examples

881. The model has been applied to: (1) perceptual skills training for air intercept control (Schneider, Vidulich, and Yeh, 1982), and (2) electronic troubleshooting training (Logan, 1988).

Technical Details

882. Again, only general guidelines for instructional strategy development can be derived with this approach.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A prepared system and its control processes. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation (Vol. 2). New York: Academic Press.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful process in memory. Journal of Experimental Psychology: General, 108, 356-388.

LaBerge, D. & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. Human Factors, 30 (5), 583-598.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. Psychological Review, 84, 1-66.

Schneider, W., Vidulich, M., & Yeh, Y. (1982). Training spatial skills for air-traffic control. Proceedings of the 26th Annual Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society.

Future Needs

883. This model requires: (1) a series of transfer-of-training studies, part-task to whole task, to demonstrate its practical utility and generalizability; and (2) assessment of its potential use for prescriptive purposes (i.e., to derive training guidelines).

7.4 MATHEMATICAL MODEL SUMMARIES

7.4.1 Learning Curve Models

Summary Description

884. Learning curve models are mathematical models that describe the temporal aspect of improved performance over time. Learning curve models involve the fitting of curves to data in order to reveal important aspects of skill acquisition.

History and Source

885. T.P. Wright initiated the use of learning curve modelling to predict production rates in the airframe industry. Since that time, modelling of learning curves has been directed toward establishing wage incentive plans, comparing jobs on difficulty, personnel selection, determining when to terminate training, and understanding and improving the process of learning. Most of the research sought to validate specific postulates of various learning theories (Spears, 1985). A few researchers have used learning curve models to explain the influences of experimental variables, or to simply identify behavioural examples of learned concepts (e.g., Baird and Noma,1978; Restle and Greeno, 1970).

Product and Purpose

886. Learning curve models offer insight into: how pre-training estimates may impact the rate of learning, when learning occurs, the predictive power of early learning on proficiency, and the points at which rapid learning (snowballing) ceases and refinement (honing) of learned skill begins.

When Used

887. Use with processes that involve a high level of operational control. If data is generated during training, these formulas can guide training pace and focus. When control group comparisons are not feasible, these models may indicate the impact of training variables.

Procedures for Use

888. Learning curve models assume methods exist for recording: number of errors made during training, or expert ratings of trainee proficiency on a number of tasks, or mean expert ratings across conditions. Also assumed is the existence of terminal trainee performance ratings. "Terminal" refers to data collected some time after training, as ratings collected immediately after training may not allow sufficient time for any cognitive assimilation required for performance. Data collected is entered into the learning curve model and results are plotted for visual inspection. Once a model is validated, individual performance assessments and predictions can be made.

Advantages

889. The use of "constants" from curve fitting (e.g., asymptotic or beginning level, rate of change) measure variables of interest more reliably than measures directly obtained. This is because the "constants" are based on entire data patterns rather than particular individual

observations within a data set. Curve fitting also smooths out random irregularities in data patterns. The constants can provide measures of learning and transfer that are not available by means commonly in use, and can be used for data from individual subjects.

Limitations

890. According to Johnson (1985) the modelling of learning data has been directed more toward the goal of understanding and improving the process of learning, rather than quantitatively describing the process for predictive purposes. This has often taken the form of flow charts and block diagrams rather than mathematical equations.

Application Examples

891. Learning curves have been applied to many different situations: (1) to compare the difficulty of performing different jobs (Dudley, 1968); (2) in personnel selection and prediction of job success (Glover, 1966; Sriyananda and Towill, 1973); and (3) to establish the performance level at which training should be terminated (i.e., training criteria) (Knowles and Bell, 1950).

Technical Details

892. Spears (1985) demonstrated the use of various constants derived by fitting equations to training and performance data. Towill (1989) discussed the pros and cons of different learning curve equations and made recommendations as to when and why each should be used.

References

Baird, J. C., & Noma, E. (1978). Fundamentals of scaling and psychophysics. New York: Wiley.

Dudley, W. A. (1968). Work management: Some research studies. London: Macmillan.

Glover, J. H. (1966). Selection of trainees and control of their progress. International Journal of Production Research, 5 (1). 43-60.

Johnson, S. L. (1985). Using mathematical models of the learning curve in training system design. Proceedings of the Human Factors Society 29th Annual Meeting, 2, 735-766. Santa Monica, CA: Human Factors Society.

Knowles, A. R. & Bell, L. F. (1950). Learning curves which tell you who's worth training and who isn't. Factory Management and Maintenance, 103, 114-115.

Restle, F. & Greeno, J. G., (1970). Introduction to mathematical psychology. Reading, MA: Addison-Wesley.

Spears, W. D. (1985). Measurement of learning and transfer through curve fitting. Human Factors, 27 (3), 251-266.

Sriyananda, H. & Towill, D. R. (1973). Prediction of human operator performance. IEEE Transactions on Reliability, R22, 148-156.

Towill, D. R. (1989). Selecting learning curve models for human operator performance. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton, & L. Van Breda (Eds.), Applications of human performance models to system design. New York: Plenum Press.

Future Needs

893. Models needn't be used only to demonstrate theoretical postulates. Applied researchers need to describe actual behaviour and transfer of training by fitting learning curve models to collected data. Assess the utility of combining the four learning curve models reviewed by Spears (1985): asymptotic, beginning level, rate constant, and inflection of learning in one data base.

7.5 TASK-BASED MODEL SUMMARIES

7.5.1 Modeling of Armor Procedures

Summary Description

894. Sticha, Knerr and associates (1985) have developed models of learning and performance for eight military armor procedural tasks. The models developed combine a network representation (using the SAINT simulation system, see Chapter 8) of performance for the eight tasks, with psychological models of skill acquisition and retention. Maximum likelihood estimates can be derived and model predictions can be compared to empirical data for validation.

895. The psychological models describing learning, retention, and recall of individual task elements are used as subroutines within the SAINT models. Then, these subroutines interact with SAINT user-defined task characteristics that represent conditions (e.g., strength) of memory for the task. This representation allows the calculation of several parameter values on various measures of performance. Skill acquisition parameters can be assigned independently for each element of a procedure.

History and Source

896. This modelling approach evolved from the merging of two components. The first one is the SAINT system, which is described in Chapter 8 of this report. The second component is the psychological models. These models that describe learning and retention are based on the concept of the strength of an association (Wickelgren, 1974). Acquisition is described by a function relating association strength to the amount of practice, or number of training trials. The function incorporated in the models also follows the Hullian assumptions (Hull, 1943).

Product and Purpose

897. With the aid of the simulation software, the user can calculate and display the proportion of correct responses on each task element, across all tasks, and the performance time. The SAINT models also provide a graph of performance by trial.

When Used

898. As a validation tool for existing training programmes that can be broken down into linked task elements performed within one larger team exercise.

Procedures for Use

899. Write performance criteria for individual task performance (subroutines), test individuals and track performance on individual task performance, develop estimates of performance, map empirically observed performance against expected (estimated) performance. Retest individuals at a later time to assess any decay of performance.

Advantages

900. There are four main advantages to this type of modelling:

(1) These models have been validated with empirical data. The models were shown to predict accurately: (a) overall performance improvements during training; and (b) the decay in performance shortly after training.

(2) Since the psychological models can be separated from the simulation, a more rigorous model validation can be conducted.

(3) Models can be used as a decision support system for training managers.

(4) Models can be used to organise the results of learning and retention experiments, and guide the researcher for future applications.

Limitations

901. These models are not intended for initial training design and development or small (individual) training evaluations. They assume each team member's task functions are defined and errors are easily identified.

Application Examples

902. Knerr and Sticha have used the model for assessment of eight military armor procedural tasks.

Technical Details

903. Maximum likelihood estimates must be calculated using commercially available mainframe-based statistical packages, and their fit with the collected data produces an empirical validation of training effectiveness. Parameters of skill learning can then be calculated for each individual task performed.

References

Hull, C. L. (1943). Principles of behaviour. New York: Appleton.

Knerr, C. M. & Sticha, P. J. (1985). Models of learning and performance of armor skills. Proceedings of the Symposium on the Military Value and Cost Effectiveness of Training (491-512), Brussels, Belgium: NATO.

Wickelgren, W. A. (1974). How to solve problems. San Francisco: Freeman.

Future Needs

904. According to Knerr and Sticha, the most pressing need is research and development on estimating the values of model parameters without collecting considerable learning and retention data.

7.6 SYSTEM-ORIENTED MODEL SUMMARIES

7.6.1 Optimization of Simulation-Based Training Systems (OSBATS)

Summary Description

905. This model is designed to allow the developers of training devices to ask "what if" questions about training device requirements and alternative solutions for those requirements. The basic goal is to specify methods to develop training concepts, training device design, and allocations of time to different training alternatives that either minimise the cost required to meet training objectives, or maximise the training effectiveness obtained at a specified cost.

History and Source

906. The model is based on several lines of research: (1) work on the optimal allocation of training time between simulators and operational equipment; (2) empirically based prescriptions for training-system design; (3) mathematical representation of skill acquisition, retention, and transfer; (4) analyses of factors affecting simulator cost and training effectiveness; and (5) simulator fidelity.

Product and Purpose

907. The OSBATS models have been developed to date by taking a top-down analytical approach. The overall problem of training device design was decomposed and five problem areas have been addressed based on the task, equipment, and training variables involved in training device design and use. In this way a modelling framework was developed that allows the addition and insertion of new models (referred to as modules) for different aspects of the training device concept formulation process. Each of the five areas is addressed through the development of a different module. The modules currently implemented are:

(1) Simulator Configuration

(2) Instructional Feature Selection

(3) Fidelity Optimization

(4) Medium Selection

(5) Resource Allocation

908. The simulation configuration module involves the selection of one of several classes of training devices, based on the task and its training requirements. The instructional feature selection module uses task training requirements and the cost of training on actual equipment to determine the feature mix by applying a set of rules for feature selection. The fidelity optimisation module specifies the realism (fidelity) required in the training device by matching different levels of various fidelity dimensions in order to get an estimate of the benefit based on transfer to the actual equipment. The media selection module aids in selecting the best training media for each task, or group of tasks. It also produces criterion estimates for training the task

on the media, estimates the overall life cycle costs of the training device, and produces an estimate of the training effectiveness for the training device. Finally, the resource allocation module produces a chronological sequence and set of training times for the training devices.

Procedures for Use

909. Due to the complexity of OSBATS, it cannot be applied "as is." There are needs for the development of software, scaling procedures and model expansions if validation is required. The model is intended to an "ideal" method of training system design.

Advantages

910. The OSBATS models are meant primarily for use by professionals (i.e., school personnel, engineers, and contractors) involved in training device concept formulation and design efforts. The OSBATS model was developed for implementation on a computer, so that the user can rapidly exercise the models. In this way the different task sets, different cost considerations, and varying instructional approaches can be tested for their effect on the training device configurations and projected effectiveness of the training systems. In addition, the system provides an audit trail of the information used and the decisions made during the use of the system.

Limitations

911. The research base for developing the models has numerous gaps. Research topics which need to be more fully addressed include: the development of task-analytic methods for estimating learning and retention parameters, the development of methods for predicting transfer between tasks/courses, development of a psychological fidelity model, and development of methods to predict media costs. In addition, the data sets which are required to allow the model to fully function are limited in many cases.

Application Examples

912. The model has been applied to the Cobra helicopter simulator and an armor maintenance job.

Technical Details

913. OSBATS addresses five common training design areas in a top-down approacn. The model is based on research in the areas of: simulator fidelity, simulator cost and training effectiveness, mathematical models of skill acquisition, empirical training-system design, and optimal use of training versus operational equipment. An interactive prototype of the OSBATS software has been successfully tested which runs on IBM hardware. Sticha (1989) described the software's approach to the five development areas.

References

Sticha, P. J. (1989). Normative and descriptive models for training system design. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton, & L. Van Breda (Eds.), Applications of human performance models to system design. New York: Plenum Press.

Sticha, P. J., Blacksten, H. R., Mumaw, R. J., Morrison, J. E., DeYoe, P. W., Cross, K. D., Buede, D. M., & Zirk, D. A. (1986). Optimization of simulation-based training systems Vol. 1. Plans for model implementation, evaluation, and research (Final Report). Alexandria, VA: Human Resources Research Organization.

Future Needs

914. Address research gaps to better support the model. Identify applications where the full model can be realised and refined.

7.6.2 Automated Simulator Test and Assessment Routine (ASTAR)

Summary and Description

915. This model is an analytical technique for predicting and assessing the effectiveness of training devices. Its greatest potential application is during the design phase, when alternative training device configurations can be evaluated to determine their relative predicted effectiveness.

History and Source

916. ASTAR builds upon prior research into analytical effectiveness prediction techniques, notably the TRAINVICE and DEFT projects (Rose, Evans and Wheaton, 1987; Rose and Martin, 1984).

Product and Purpose.

917. The ASTAR system is a set of computer programmes which enables the entry of data on a training device and the trainees, and provide outputs which assess the training effectiveness of the device. The programmes are written for the IBM PC series of computers.

When Used

918. The ASTAR system is intended for use during the design of a training device, although it may also be used to assess existing training devices.

Procedures for Use

919. The user enters a list of tasks to be trained into the ASTAR system. ASTAR then asks a series of questions concerning each task. These questions are answered, on line, by a person who has the necessary knowledge of the training device design, the trainees, and the training situation. ASTAR then operates on this data to generate a prediction of device effectiveness. ASTAR also provides diagnostic information which helps to isolate problem areas in the device.

Advantages

920. ASTAR allows device effectiveness evaluation without empirical studies, which are expensive and sometimes not feasible.

Limitations

921. The validity of the ASTAR predictions has not been fully assessed. Furthermore, the ASTAR analysis requires the availability of an appropriate task listing, and participation of persons with the necessary knowledge of the device, trainees, and training situation.

Application Examples

922. During its development the ASTAR system was used to assess an anti-submarine tactics trainer, an avionics maintenance trainer, and gunnery trainer.

Technical Details

923. ASTAR is an interactive, menu driven computer programme written in COBOL designed to be used on an IBM (or compatible) personal computer. ASTAR divides the evaluation of a training device into four areas: training problem analysis, training efficiency analysis, transfer problem analysis, and transfer efficiency analysis. The questions asked by ASTAR and the algorithm by which it determines predicted effectiveness are based on existing training research (Knerr, Sticha, and Blacksten, 1989).

References

Knerr, C. M., Sticha, P. J., & Blacksten, H. R. (1989). Human performance models for training design. In G. R. McMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton, & L. Van Breda (Eds.), Applications of human performance models to system design. New York: Plenum Press.

Rose, A. M., Evans, R., & Wheaton, G. R. (1987). Methodological approaches for simulator evaluations. In S. M. Cormier & J. D. Hagman (Eds.), Transfer of training: Contemporary research and applications. San Diego: Academic Press, Inc.

Rose, A. M. & Martin, A. W. (1984). Forecasting device effectiveness: III. Analytic assessment of DEFT (Contract No. MDA903-82-0414). Washington, D.C.: American Institute for Research.

Rose, A. M., Martin, A. W., & Yates, L. G. (1985). Forecasting device effectiveness: III. Analytic assessment of device effectiveness forecasting technique (Technical Report 681). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Rose, A. M. & Wheaton, G. R. (1984). Forecasting device effectiveness: II. Procedures. Washington, D.C.: American Institute for Research.

Rose, A. M., Wheaton, G. R. & Yates, L. G. (1985). Forecasting device effectiveness: I. Issues (Technical Report 680). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Future Needs

924. Further studies of the validity of the ASTAR predictions are required.

This page has been left blank intentionally.

CHAPTER 8

NETWORK MODELLING TOOLS

8.1 INTRODUCTION

925. In many discussions of human operator models, there is often confusion between what might be called a human operator model such as HOS and an operator modelling technology such as SAINT. It is important that a user, such as a system design engineer, understand the difference so that he can make an informed choice and, hopefully, not be disappointed.

926. A human performance model describes, and/or addresses, some specific subset of human performance issues or actions. In this sense then, a model deals with a portion of the potential issues handled by modelling tools. Models come equipped with either data or the templates to store the data. The data formats and requirements are often established within the structure of the model. The input, the internal processing and calculations, and the output are predefined. Often, a model contains specific mathematical algorithms to predict or estimate some aspect of human performance. However, models are generally not specific to a particular system and the system, tasks, number and type of operators are the variables which provide a model with its flexibility and utility. Nonetheless, the algorithms, and the models themselves, are generally theory based or derived from established human performance principles. Since the model is intended to be an accurate representation of human behaviour it is essential that it be valid. Validity is critical.

927. A modelling tool, on the other hand, is a set of capabilities for modelling the human operator. It comes void of the type of structure incorporated into models, although it contains capabilities to represent these specifics. It is designed for use in a problem solving situation and includes only those human performance parameters of interest to the problem solver. A modelling technology is judged more on its utility than its validity. If the modelling technology provides the problem solver (e.g., design engineer) with even a gross estimate of how well an operator will perform in a given design configuration and does so in a manner that is relatively easy to obtain, then the modelling technology may be declared successful in having done its job for the user. One still feels uneasy about ignoring validity but the utility brings with it a kind of "validity by acclamation".

928. The advantage of the human operator model is that all that is required is to provide the input conditions and it should provide a solution. The disadvantage comes in a limited parameter set which may not be relevant to the problem at hand and the very detailed human performance output which may be much more elemental than desired. The advantage of the modelling technology is that one need only model the problem at hand. The user is provided with some very powerful techniques for describing the various tasks and their linkage. The disadvantage is that the user must provide the specifics of the system structure.

929. The authors believe that in terms of aiding the design engineer, the benefits are more likely to come from modelling technologies rather than actual models of operator performance. The latter models tend to be very complex because they are dealing with a very complex entity (i.e., Homo Sapiens). They generate output which may be overly specific to the user need. However, these models have been of tremendous importance in highlighting the major human performance variables which must be addressed by the user. Also, experience in

developing human operator models has had an extremely valuable heuristic benefit in the development of modelling technologies. The modelling technologies to be addressed in this chapter include SAINT, Micro-SAINT, and SLAM.

8.2 OVERVIEW AND RECOMMENDED REFERENCES

930. As noted in the introduction, modelling technologies appear to be on the verge of becoming an accepted design tool by the system design engineer. Among the models reviewed in this chapter, the oldest and pioneer model in the field is SAINT. SAINT's capability to model both continuous and discrete operations made it a very attractive model for system designers. However, the complexity of the model and accompanying software support requirements made it a capability that only a few large system developers could afford. But few people desired the large number of applications which could be realized by a model such as SAINT.

931. There have been two major outgrowths from the original SAINT modelling technology which have been developing in a concurrent fashion. One of these is MicroSAINT. The other is SLAM. MicroSAINT began as a straightforward attempt to put SAINT on a micro-computer and ended up being a much more user-friendly model than SAINT. It makes extensive use of menu driven tables to assist the user. MicroSAINT now has graphics capability for drawing task networks and describing data outputs. MicroSAINT was developed for application on the IBM PC family of micro-computers.

932. The developer of SAINT, Alan Pritsker and his company, Pritsker and Associates have developed a model called SLAM as a means of providing easier access to SAINT-like capabilities. SLAM is more suitable for a mini-computer than a micro-computer. Pritsker has added a graphics package called TESS which may be employed once the basic SLAM techniques have been acquired. Both Pritsker and Associates and Micro Analysis and Design offer training programs to learn how to use SLAM/TESS and MicroSAINT.

8.2.1 References

Meister, D. (1985). Behavioral analysis and measurement methods. New York: John Wiley & Sons.

Pew, R. W., Baron, S., Feehrer, C. E., & Miller, D. C. (1977). Critical review and analysis of performance models applicable to man-machine systems evaluation (Report AFOSR-TR-77-0520). Cambridge, MA: Bolt, Beranek and Newman, Inc., (ADA038597).

8.3 MODEL SUMMARIES

8.3.1 SAINT (Systems Analysis of Integrated Networks of Tasks)

Summary Description

933. SAINT is a FORTRAN-based, network modelling and simulation technique developed to assist in the design and analysis of complex man-machine systems. It consists of a symbol set and the capabilities to represent discrete task elements, continuous state variables, and to dynamically modify performance via moderator functions. In addition to the standard processing subroutines incorporated into the software package, SAINT also includes dummy versions of user-written subprograms which enable the analyst to tailor input, output, and simulated performance processing. The language provides extensive error checking and error messages to the user. The SAINT program enables a modeler to represent multiple operator systems and to execute multiple networks simultaneously.

934. A SAINT network consists of resources, tasks performed by resources, precedence relationships between tasks, task performance characteristics, the flow of information through the system, and the effects of environmental stressors on task performance. In SAINT, tasks are the central elements of a network. A SAINT task is characterized by parameters that specify the nature of the predecessor tasks, task characteristics and branching to other tasks. Precedence relationships specify the flow of operations through a network and the completion of individual tasks in a network can modify later precedence relationships, thereby altering network flow. Time to perform a task is specified in terms of a variety of sampling distributions such as constant, normal, lognormal, Poisson, and beta. SAINT can simulate six types of tasks: single operator, joint operator, one of several operators, hardware, cyclic tasks and gap filled tasks. After a task has been completed, SAINT decides which of the remaining tasks shall be initiated. The decision is based on five decision rules: Determination (all branches selected); Probabilistic (selection on a random basic); Conditional, Take first (first branch satisfying specified conditions is selected); Conditional, Take-all (all branches satisfying specified conditions are selected); and Modified probabilistic (same as probabilistic except branch probabilities are modified by number of previous completions of the task from which the branches stem).

History and Source

935. SAINT was developed in the 1970s by Pritsker & Associates, Inc. for the U.S. Air Force (USAF) Aerospace Medical Research Laboratory (AMRL). SAINT evolved through a series of major modifications and embellishments to the GASP and GERT modelling techniques developed during the late 60s and early 70s. SAINT I basically upgraded P-GERT to handle discrete time varying networks. SAINT II provided the capabilities to model continuous state variables and to dynamically update attributes. SAINT III incorporated moderator functions and additional attributes.

936. Documentation is available through the Defense Technical Information Center (DTIC).

Product and Purpose

937. SAINT provides the conceptual framework and the tools to develop system models in which men, machines, and the environment are represented. It enables the analyst to investigate the impact of modifications to the man-machine-environment interface on human and overall system performance. Available outputs generated by SAINT include data input reports, iteration reports, mission summary reports, and user generated reports. These reports provide information concerning resource utilization, task performance, state variable status, system performance measures, etc. Output information can be obtained in the form of summary tables, statistical plots, and/or histograms.

When Used

938. SAINT was designed to assess and analyze human performance in manned systems. However, it is applicable to any dynamic, time varying system in which discrete and continuous elements are to be modelled and simulated. In addition, SAINT should be used when the requirement calls more for a general purpose computer language than a model per se. The software package can be used to model and evaluate systems at any stage of development.

Procedures for Use

939. In order to use SAINT, the user must first develop a network model of the system using the SAINT symbol set. Then, the network model is converted into data cards/records readable by the simulation program's processor. The input to the SAINT program is an alphanumeric representation of the model. The organization and content of the input deck describes the tasks, task data, resources, network structure, etc. If nonstandard processing is needed (i.e., tailored input, output, or dynamic modification of task performance), then the user must include the lines of FORTRAN instruction into the appropriate SAINT user subprogram. Once the input deck is prepared, it is submitted to the program for processing. The execution of the program simulates the model, calculates estimates of operator and system performance, and outputs the statistical descriptions of the performance. The user then analyses and assesses the output. Detailed procedures for using SAINT are contained in the SAINT Users Manual (Wortman, et. al, 1977, 1978).

Advantages

940. SAINT is one of the earliest languages which permitted the user to represent multiple operator systems and to execute multiple networks simultaneously. Another major advantage is its ability to model human tasks either as discrete or continuous processes. In sum, it is a highly flexible tool which may be used in a variety of applications. Also, the language is the property of the USAF and, therefore, available to U.S. Government agencies and companies contracted to the government.

Limitations

941. While SAINT does have an elementary level requiring no programming experience, exploitation of much of the SAINT sophistication requires programming expertise. There is a significant amount of software support associated with using SAINT.

Application Examples

942. SAINT has been used to develop models of a choice reaction time task, a remotely piloted vehicle control facility, a Digital Avionics Information System display, an airborne warning and control system, the performance of industrial inspectors, and the AN/TSQ-73 air defense command and control system.

Technical Details

943. The SAINT software is a FORTRAN-based language and therefore transportable across a variety of systems. The random number generator contained in the program is system specific and must be tailored to the host computer. The language consists of approximately 12,000 lines of instruction and requires about 325K bytes of memory. The program is available for Digital Equipment Corporation VAX 11/700 series mini-computers and IBM mainframe computers. The SAINT software is also commercially available for the Apple Macintosh and the IBM PC micro-computers.

References

Duket, S. D., Wortman, D. B., Seifert, D. J., Hann, R. L., & Chubb, G. P. (1978). Documentation for the SAINT simulation program (AMRL-TR-77-63).Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory (AD A059198).

Pritsker, A. A. B., Wortman, D. B., Seume, C. S., Chubb, G. P., & Seifert, D. J. (1974). SAINT: Vol. I. Systems analysis of an integrated network of tasks (AMRL-TR-73-126). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory.

Seifert, D. J., Koeplinger, G., & Hoyland, C. M. (1980). REDIMEN: SAINT Redimensioning Program. (AFAMRL-TR-80-5). Wright-Patterson AFB,OH: Aerospace Medical Research Laboratory.

Wortman, D. B., Duket, S. D., & Seifert, D. J. (1976). SAINT simulation of a remotely piloted vehicle/drone control facility: Model development and analysis (AMRL-TR-75-118). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory.

Wortman, D. B., Duket, S. D. Seifert, D. J., Hann, R. L., & Chubb, G. P. (1978). Simulation using SAINT: A user-oriented instructional manual. (AMRL-TR-77-61). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory, (AD A058671).

Wortman, D. B. Duket, S. D., Seifert, D. J., Hann, R. L., & Chubb, G. P. (1977). The SAINT user's manual (AMRL-TR-77-62). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory

Wortman, D. B., Hickson, E. F. III, & Jorgensen, C. J. (1978). A SAINT model of the AN/TSQ-73 guided missile air defense system. Proceedings of the Winter Simulation Conference.

Wortman, D. B., Seifert, D. J. & Duket, S. D. (1976). New developments in SAINT: The SAINT III simulation program (AMRL-TR-75-117). Wright-Patterson AFB, OH: Aerospace Medical Research Laboratory (AD A059198).

Future Needs

944. SAINT should be made more "user friendly" so that extensive programming experience is not required by the user.

8.3.2 Micro SAINT

Summary Description

945. Micro SAINT is a microcomputer version of the popular modelling language, SAINT. This version offers an easy way of entering task networks and conducting simulations of human operators in systems. It offers an interactive, menu-driven, user interface. While more limited than the full version of SAINT, it is far more accessible and easier to use. Micro SAINT consists of software to support the following five major components: (1) interactive model development; (2) interactive model execution; (3) analysis of results; (4) utilities; and (5) error messages to the user. Interactive model development entails such activities as entering and modifying tasks and task information, and defining task networks. During interactive model execution, the user can manipulate model variables, pause execution, obtain snapshots of execution, etc. The analysis of results portion of Micro SAINT enables the user to obtain statistics on the time it took to traverse the network. Utilities to copy, delete, print, and merge models are included in the software. The latest version of Micro SAINT (3.0) has a graphics capability which draws task network diagrams and data output graphics such as line, bar, and step charts, scatter plots, time lines, and frequency distributions.

History and Source

946. Micro SAINT was developed under a contract with the United States Army Medical Research and Development Command. It was conceived as a tool to evaluate the effects of pretreatment drugs on the operators of military systems. The Statement of Work for this contract called for development of a tool that would enable research scientists to simulate the effects of psychopharmacological agents on the human operators in military systems, such as tanks and helicopters. The motivation for using a simulation tool to conduct the necessary research is to avoid having to experiment with human beings.

947. Micro SAINT began as an attempt to provide the power and flexibility of the existing SAINT language on a micro computer. A complementary goal was the development of a simple and menu driven interface for the less than experienced user. It soon became obvious, however, that the complexity of psychopharmacology, coupled with the lack of solid performance data characterizing the pretreatment drugs of interest to the U.S. Army, required a far more elaborate modelling technology than Micro SAINT was able to deliver. Although not appropriate for its originally intended purpose, Micro SAINT's usefulness for simpler simulations was obvious. The company which had developed Micro SAINT, Micro Analysis and Design, realized this potential and continued the independent development of the technique.

948. The Micro SAINT software is maintained and distributed by Micro Analysis and Design, 9132 Thunderhead Drive, Boulder, Colorado 80302, (303) 442-6947.

Product and Purpose

949. The primary goal in the development of Micro SAINT was to produce a system that is easy to learn and use. This goal was accomplished by providing a menu-driven user interface, rather than a programming language. It was designed to be "simple, hot, and deep." Simple - the software must be easy enough to learn so that it will attract users right away. Hot - the system must be exciting enough to hold a user's interest. Deep - the software must keep unfolding in new levels of complexity, so that sophisticated users will continue to find more

power as they learn more about it. The three concepts were addressed as follows: First, a survey of the different types of user interfaces revealed that a menu-driven interface would be the most user-friendly. This choice satisfied the "simple" criterion. Second, since Micro SAINT is not a compiled language, the execution of models is interactive. The model builder can pause a model, examine and change some of the variables, and resume execution. This feature satisfied the "hot" criteria. Third, Micro SAINT has a parser which allows it to interpret algebraic expressions in places where a numeric value is required. The parser satisfies the "deep" criteria. Outputs from a model can be obtained on the minimum, mean, maximum, standard deviation, and the frequency distribution of the times it took to traverse the network.

When Used

950. Micro SAINT has powerful features which enable it to model a wide variety of systems in addition to man-machine systems. If a process can be described by a flowchart, it can be modelled. Micro SAINT can be used iteratively throughout a system's life.

Procedures for Use

951. Since Micro SAINT is easy to use, most of the effort in a modelling project goes into researching the process itself, rather than into writing and debugging a computer program. The first step is to draw a diagram of the task network model. Task network models are built which show the normal sequence of activities throughout an operation. Each task is represented by a box, and arrows between the boxes show the sequence from one task to another.

952. The second step is to enter the paper model into an IBM PC or compatible microcomputer. The Micro SAINT User's Guide contains a tutorial which leads a beginning user through this process. The software is entirely menu-driven and "help" screens are always available to answer questions. Entering each task is simply a matter of filling in the blanks of a task description menu. Each task in the model has several parameters, one of which is the average time it takes to perform that task. Micro SAINT enables the user to represent task times with an algebraic formula which can depend on a number of factors, including drug dosage level, fatigue, and battlefield stress. Additional task information required consists of task number distribution type and associated mean and standard deviation, decision type (single or probabilistic), if probabilistic, the probability of each branch, and the task numbers of successor tasks.

953. The third step is to execute the model with Micro SAINT's model execution program and to collect data. Execution is interactive - that is, the user can pause the simulation, change the values of variables, and resume execution. This feature is useful for performing "what-if" analyses. Data are stored in a disk file in standard ASCII format, to facilitate their importation into statistical analysis packages.

954. The final step is to analyze the data that were collected when the model was run. Micro SAINT provides some statistical functions, and the data can be graphed with a commercial package such as Lotus 1-2-3 or Symphony.

Advantages

955. Micro SAINT is a simulation tool that is as easy to use as a spreadsheet. It does not require a simulation expert to build the models, run them, and analyze the results. The manpower required to complete a simulation project is a fraction of what would be required if a traditional programming language were used.

956. Micro SAINT runs on a popular desk top microcomputer. This means that the modelling process itself can be decentralized, with each individual having more control over and confidence in the results of the simulation. Micro SAINT enables computer tools to be used by the people who have to solve the problems, not by intermediaries.

Limitations

957. Micro SAINT was developed for a microcomputer, and consequently the scale of it is micro. Models can have up to 250 tasks, and execution is probably slower than what would be expected from a main- frame computer. For example, Micro SAINT would be an inappropriate choice for building a full-scale model of a nuclear power plant. Also, the type of statistical output and the types of model enhancements are currently limited to those contained in the software package.

Application Examples

958. Micro SAINT has been used to develop models of the M60 tank firing sequence and the sequence of inspection and maintenance activities that an F-14 fighter undergoes in the course of a single day. In addition, the U.S. Army's new LHX helicopter and its cockpit design were modelled. The model simulated a combat mission that involved entrance into the enemy zone, several combat engagements, egress from the zone, and battle damage assessment. A model was also developed to determine the amount of resources required throughout a U.S. Army helicopter pilot training course as well as to model the skill acquisition of a student pilot.

Technical Details

959. Micro SAINT 3.0 runs on an IBM PC or compatible microcomputer equipped with 512K memory, and a hard disk drive or two floppy disk drives. The package includes a 200-page User's Guide, technical support, and a 30-day trial period.

References

Archer, R., Drews, C., Laughery, K. R., Dahl, S., & Hegge, F. (1986). Data on the usability of Micro SAINT. Proceedings of the NAECON 86 Conference (pp. 855-858). Dayton, OH: NAECON.

Laughery, K. R. & Drews, C. (1985). Micro SAINT: A computer simulation system designed for human factors engineers. Proceedings, Human Factors Society - 29th Annual Meeting, (1061-1064) Santa Monica, CA: Human Factors Society.

Laughery, K. R., (1984). Computer modelling of human performance on microcomputers. Proceedings of the Human Factors Society 28th Annual Meeting (pp. 884-888). Santa Monica, CA: The Human Factors Society.

Future Needs

960. The speed and power of Micro SAINT will undoubtedly improve with advances in micro-computers.

8.3.3 Simulation Language for Alternative Modeling (SLAM)

Summary Description

961. The Simulation Language for Alternative Modeling (SLAM) is an advanced, FORTRAN-based tool and the first language that provides three different modelling viewpoints in a single integrated framework. SLAM permits discrete event, continuous, and network modelling perspectives to be implemented in a single model. The SLAM language consists of network symbols, network input statements, control statements, COMMON variables, user-callable subprograms, and the basic statements for user-written subprograms. It includes the capabilities to represent multi-operator systems, to execute multiple networks simultaneously, and to model a large class of systems as networks. The network symbol set enables the modeler to build graphic models, consisting of a set of interconnected symbols, that depict the operation of the system, and that directly correspond to SLAM network input statements.

962. A SLAM network structure consists of specialized nodes and branches that are used to model resources, queues for resources, activities, and network flow decisions. A system model is represented as a set of entities (any object, being, or unit of information) which flows through a network of interconnected nodes (workstations, machinery, storage locations, etc.). Entities compete for resources (tools, machines, operators, etc.) when flowing through the system. An entity is defined, or characterized, by temporal and physical feature information stored in its "attribute" array. The flow of entities normally follows the directed branches indicated on the network and this flow results in changes in the state of the simulated system. Branches, which can be either probabilistic or deterministic, are used to depict the passage of time and activities performed in relation to entities

963. The SLAM software also includes dummy versions of subprograms which can be user-written by the analyst to perform non-standard network processing, specialized output reporting, continuous variable definition, and discrete event scheduling. The software library also contains 11 probability distribution functions (e.g., uniform, triangular, exponential, normal, Poisson), and a user-definable function, which can be used for sampling arrival times, performance times, production rates, etc Finally, the language contains approximately 50 execution error messages to alert the modeler to problems.

History and Source

964. SLAM was developed by Pritsker & Associates (P&A) by combining and evolving the GASP IV and Q-GERT simulation languages into a single integrated framework. It was introduced by P&A in 1979 and, in 1981, significant improvements were made to the internal operations of SLAM, resulting in SLAM II. SLAM II has been updated several times, with the latest enhancements being in Version 2.3 (1984) or a higher version. Improvements include simplified model design, decreased execution time, and expanded statistical outputs. In two other developmental efforts by P&A, SLAM II has been adapted to run on a microcomputer and the need for a graphics capability was addressed in a program called TESS (The Extended Simulation System). TESS enabled users to build SLAM networks graphically at a terminal and to describe system operation through animation. Once the graphic network has been completed, the TESS program converts the graphic symbols to input statements for the SLAM processor.

965. The SLAM program is maintained and distributed by Pritsker & Associates, Inc., P.O. Box 2413, West Lafayette, Indiana 47906, from whom copies of the source tape may be purchased. Basic documentation of the SLAM II modelling methodology is provided in Pritsker (1986).

Product and Purpose

966. SLAM provides a set of concepts, procedures, and techniques to represent the dynamic behaviour of any system. The language combines network, discrete event, and continuous modelling capabilities into a unified modelling framework. It can be used for system design analyses, procedural analyses, and performance assessments at any stage of system development. Output reports generated by SLAM include the input listing and error messages, echo report, trace report, and summary report. The summary report includes the statistical results from the execution of a simulation model and consists of statistics for queues, files, activities, gates, discrete and continuous variables, and resource utilization. Statistical information includes values for the mean, minimum, maximum, standard deviation, gate status (open/closed), resource capacity, etc. The analyst can also specify and tailor the type of statistics to be collected and output by means of 29 statistical calculation functions and the various report writing subroutines.

When Used

967. SLAM can be used during any phase of a developmental or existing system's life. Virtually any system of interest can be modelled using SLAM II, including production lines, transportation networks, communication networks, and complex military man-machine weapons systems. The tool can be employed for analyzing a problem situation, for specific decision making, for designing a new system, for redesigning an existing system, or for projecting future developments concerning a system.

Procedures for Use

968. The essential task for the analyst is to utilize the SLAM network concepts to formulate a network model which reflects the important characteristics of the system. First, the analyst must state the problem that the simulation model is to address. Then one must define the elements of the system which are to be represented as entities. Next, one must construct the network of nodes and activities through which the entities flow. This involves the preparation of network input statements, and the sequencing of these statements. Then, the network description statements are combined with the necessary control statements into a computer file. The modeler must also write any required FORTRAN statements for the user-defined subprograms to be used in the simulation. All model relevant FORTRAN and SLAM files are then compiled, linked, and the model executed. Finally, the analyst evaluates the simulation output.

Advantages

969. SLAM provides a flexible and portable language which runs on a wide variety of computing systems, including mainframes as well as personal computers. The program has been installed in more than 400 industrial, academic and government installations, and has contributed to the increased use of modelling and simulation throughout the world.

Limitations

970. SLAM capabilities are oriented primarily towards network--process control--models and, therefore, the modelling of man/machine systems becomes somewhat more cumbersome in SLAM than in SAINT. The networking capabilities feature the flow of items to service stations whereas man/ machine systems are often represented as the flow (movement) of operators through the performance of a task sequence. This difference in perspective renders some of the output from the built-in statistical routines meaningless. Also, the use of the discrete event routines, in essence, requires the user/analyst to write the necessary FORTRAN statements. In addition, SLAM is a complex language consisting of numerous types of nodes, input statements, control statements, subroutines, functions, and variables. This complexity makes it necessary for the potential SLAM modeler to obtain formal training at P&A facilities and to invest a sizable amount of time and experience in its use. The SLAM software is proprietary to P&A and must be purchased from them.

Application Examples

971. SLAM II has been used to model the inspection and adjustment stations on a production line, an inventory system with lost sales and back orders, the servicing of customers at drive-in bank windows, and the operator engagement sequence for U.S. Army short range air defense systems. Many of these applications are described in Pritsker (1986).

Technical Details

972. SLAM is a generalization of the GASP and GERT languages which resulted from the combination of the networking features of Q-GERT and the discrete and continuous modelling capabilities of GASP IV. It is a FORTRAN-based language consisting of approximately 13,000 lines of code. The software program is available for mainframe computers, minicomputers and the IBM Personal Computer.

References

Pritsker, A. A. B. (1986). Introduction to simulation and SLAM II (3rd ed.). New York: John Wiley & Sons, Inc.

O'Reilly, J. J. (1984). SLAM II quick reference manual. Pritsker & Associates, Inc.

Future Needs

973. No future modifications to the program are discernible at the present time.

This page has been left blank intentionally.

LIST OF PARTICIPANTS

AUTHORS OF THIS REPORT

Chapter 1 - Dr. Grant R. McMillan
Human Engineering Division
Armstrong Laboratory
Wright-Patterson Air Force Base, OH 45433-6573
U.S.A.

Chapter 2 - Mr. David Beevis
Human Engineering Section
DCIEM
P.O. Box 2000
Downsview, Ontario, M3M 3B9
Canada

Chapter 3 - Dr.-Ing. Willi Stein
Forschungsinstitut fur Anthropotechnik
(FGAN/FAT)
Konigstrasse 2
D-5307 Wachtberg-Werthhoven
Federal Republic of Germany

LCDR Robert Sutton
Royal Naval Engineering College
Manadon, Plymouth, Devon, PL5 3AQ
United Kingdom

Chapter 4 - Dr.-Ing. Willi Stein
Forschungsinstitut fur Anthropotechnik
(FGAN/FAT)
Konigstrasse 2
D-5307 Wachtberg-Werthhoven
Federal Republic of Germany

LCDR Robert Sutton
Royal Naval Engineering College
Manadon, Plymouth, Devon, PL5 3AQ
United Kingdom

Mr. David Beevis
Human Engineering Section
DCIEM
P.O. Box 2000
Downsview, Ontario, M3M 3B9
Canada

Chapter 5 - Dr. Michael H. Strub
Army Research Institute Field Unit
P.O. Box 6057
Fort Bliss, TX 79906-6057
U.S.A.

Mr. Kenneth C. Reynolds
Army Research Institute Field Unit
P.O. Box 6057
Fort Bliss, TX 79906-6057
U.S.A.

Chapter 6 - Mr. David Beevis
Human Engineering Section
DCIEM
P.O. Box 2000
Downsview, Ontario, M3M 3B9
Canada

Contributor to subsections 6.6.2 and 6.6.5:
Dr. Joe W. McDaniel
Human Engineering Division
Armstrong Laboratory
Wright-Patterson Air Force Base, OH 45433-6573
U.S.A.

Chapter 7 - Dr. Eduardo Salas
Human Factors Division (Code 262)
Naval Training Systems Center
Orlando, FL 32813-7100
U.S.A.

Chapter 8 - Dr. Michael H. Strub
Army Research Institute Field Unit
P.O. Box 6057
Fort Bliss, TX 79906-6057
U.S.A.

Mr. Kenneth C. Reynolds
Army Research Institute Field Unit
P.O. Box 6057
Fort Bliss, TX 79906-6057
U.S.A.

CHAIRMEN OF RSG.9

Prof. Dr. G. Johannsen, F.R. Germany (1983-1984)

Mr. P.M. Linton, U.S.A. (1984-1985)

Dr. G.R. McMillan, U.S.A. (1985-1991)

MEMBERS OF RSG.9

Canada

Mr. D. Beevis (1983-1991)

France

Prof. Dr. A. Coblentz (1983-1991)

F.R. Germany

Prof. Dr. G. Johannsen (1983-1984)
Dr.-Ing. W. Stein (1983-1991) - Secretary of RSG.9

Netherlands

Dr. P.H. Hudson (1983)
Prof. Dr. W.A. Wagenaar (1984-1985)
Prof. J. Moraal (1985-1987)
Mr. L. van Breda (1987-1991)

United States

U.S. Army
Mr. S.R. Stewart (1983-1984)
Dr. M.H. Strub (1984-1991)

U.S. Navy
Mr. P.M. Linton (1983-1985)
Dr. N.E. Lane (1984)
Dr. R.E. Perryman (1985)
Dr. E. Salas (1986-1991)

U.S. Air Force
Dr. G. McMillan (1983-1991)

United Kingdom

Mr. C. Shepard (1983-1984)
Dr. R. Sutton (1985-1991)

DRG DOCUMENT CENTRES

NATO does not hold stocks of DRG publications for general distribution. NATO initiates distribution of all DRG documents from its Central Registry. Nations then send the documents through their national NATO registries, sub-registries, and control points. One may sometimes obtain additional copies from these registries. The DRG Document Centres listed below can supply copies of previously issued technical DRG publications upon request.

BELGIUM
EGM-JSRL
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles
Tel:(02)243 3163, Fax:(02)243 3655

CANADA
Directorate of Scientific Information Services
National Defence Headquarters
MGen. George R. Pearkes Building
Ottawa, Ontario, K1A OK2
Tel:(613)992-2263, Fax:(613)996-0392

DENMARK
Forsvarets Forskningstjeneste
Ved Idrætsparken 4
2100 København Ø
Tel:3927 8888 + 5660,
Fax:3543 1086

FRANCE
CEDOCAR
00460 Armées
Tel:(1)4552 4500, Fax:(1)4552 4574

GERMANY
DOKFIZBw
Friedrich-Ebert-Allee 34
5300 Bonn 1
Tel: (0228)233091, Fax:(0228)125357

GREECE
National Defence Headquarters
R+T Section (D3)
15561 Holargos, Athens
Tel: (01)64 29 008

ITALY
MOD Italy
SEGREDIFESA IV Reparto PF.RS
Via XX Settembre, 123/A
00100 Roma
Tel:(06)735 3339, Fax:(06)481 4264

THE NETHERLANDS
TDCK
P.O. Box 90701
2509 LS Den Haag
Tel:(070)3166394, Fax:(070)3166202

NORWAY
Norwegian Defence Research Establishment
Central Registry
P.O. Box 25
2007 Kjeller
Tel:(06)80 71 41 Fax:(06)80 71 15

PORTUGAL
Direcção-General de Armamento
Ministério da Defesa Nacional
Avenida da Ilha da Madeira
1499 Lisboa
Tel:(01)610001 ext.4425, Fax:(01)611970

SPAIN
Ministerio de Defensa,DGAM
SDG TECIN, C/ Arturo Soria 289
28033 Madrid
Tel:(91)3020640, Fax (91)3028047

TURKEY
Genelkurmay Genel Plan Prensipler
Savunma Arastirma Daire Baskanligi
Ankara
Tel:(4)1176100 ext.1396, Fax:(4)1250813

UNITED KINGDOM
DRIC.
Kentigern House, 65 Brown Street
Glasgow G2 8EX
Tel:(041)224 2435, Fax:(041)224 2145

UNITED STATES
DTIC
Cameron Station
Alexandria, VA 22304-6145
Tel:(202)274-7633, Fax:(202)274-5280

DEFENCE RESEARCH SECTION
NATO HEADQUARTERS
B 1110 BRUSSELS
BELGIUM
Telephone [32](2)728 4285 - Telefax [32](2)728 4103
(not a DRG Document Distribution Centre)